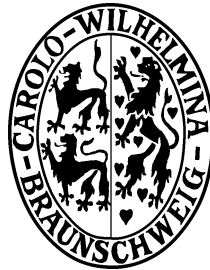


# Nichtlineare, diskrete Filteralgorithmen zur numerischen Lösung hyperbolischer Erhaltungsgleichungen



Vom Fachbereich für Mathematik und Informatik  
der Technischen Universität Braunschweig  
genehmigte Dissertation zur Erlangung des Grades einer  
Doktorin der Naturwissenschaften  
(Dr. rer. nat.)

von  
**Andrea Bürgel**

Eingereicht am: 16. September 2004

Tag der mündlichen Prüfung: 10. Februar 2005

1. Referent: Prof. Dr. Thomas Sonar, TU Braunschweig
2. Referent: Prof. Dr. Joachim Weickert, Universität des Saarlandes



# Vorwort

Die vorliegende Arbeit entstand während meiner Tätigkeit als wissenschaftliche Mitarbeiterin am Institut *Computational Mathematics* (ehemals Institut für Analysis) der Technischen Universität Braunschweig. Für die Unterstützungen und Ermutigungen, die ich bei der Entstehung dieser Arbeit von vielen Seiten erfahren habe, möchte ich mich an dieser Stelle bedanken.

Mein besonderer Dank gilt meinem Betreuer Professor Dr. Thomas Sonar, der mit seinen fachlichen Anregungen und persönlichen Ermunterungen wesentlich zum Entstehen dieser Arbeit beigetragen hat. Für seine ständige Diskussionsbereitschaft, seine fachliche Unterstützung und die gute Zusammenarbeit möchte ich ihm herzlich danken.

Herrn Professor Dr. Joachim Weickert danke ich für die freundliche Übernahme des Korreferates und für seine Einladung zu einem Aufenthalt an der Universität des Saarlandes, bei dem ich viele neue Ideen für meine Arbeit erhalten habe.

Thomas Brox aus der Arbeitsgruppe von Herrn Professor Dr. Joachim Weickert danke ich für die erfreuliche Zusammenarbeit auf dem Gebiet der vollständig diskreten Filter. Von ihm stammt die grundlegende Idee der Restflächenverteilung für den ersten, in Kapitel 4.3 beschriebenen, diskreten Filter.

Allen meinen Kollegen der Arbeitsgruppe Partielle Differentialgleichungen, insbesondere der Gruppe von Herrn Sonar, danke ich für die angenehme Zeit am Institut und für die Hilfsbereitschaft und Unterstützung, die mir während meiner Arbeit von allen Seiten entgegen gebracht wurde.

Außerdem möchte ich mich bei meinem Freund Ulli bedanken, der es sicher nicht immer leicht mit mir hatte, wenn die Arbeit nicht zu meiner Zufriedenheit voranging. Für seine Geduld und seine Unterstützung, auf die ich mich jederzeit verlassen konnte, danke ich ihm herzlich.

Andrea Bürgel



# Inhaltsverzeichnis

<b>Einleitung</b>	<b>1</b>
<b>1 Grundlagen</b>	<b>5</b>
1.1 Hyperbolische Erhaltungsgleichungen . . . . .	5
1.1.1 Skalare Erhaltungsgleichungen . . . . .	7
1.1.2 Systeme von Erhaltungsgleichungen . . . . .	12
1.2 Numerische Methoden . . . . .	13
1.2.1 Grundlegende Diskretisierungen und Eigenschaften . . . . .	14
1.2.2 Verschiedene Klassen numerischer Verfahren . . . . .	20
1.3 Testprobleme . . . . .	23
1.3.1 Eindimensionale Testfälle . . . . .	24
1.3.2 Zweidimensionale Testfälle . . . . .	28
<b>2 Grundlegende Herleitungen nichtlinearer Filter</b>	<b>33</b>
2.1 Methoden der Bildverarbeitung . . . . .	33
2.1.1 Der Variationsansatz . . . . .	34
2.1.2 Ansätze basierend auf partiellen Differentialgleichungen . . . . .	37
2.2 Vollständig diskrete Filterroutinen . . . . .	42
2.2.1 Erste diskrete Filteralgorithmen von Engquist, Lötstedt und Sjögreen . . . . .	43
2.2.2 Zusammenhang zu Methoden der Bildverarbeitung . . . . .	47
<b>3 Nichtlineare, diskrete Filter der Bildverarbeitung</b>	<b>51</b>
3.1 Digitaler TV-Filter von Chan, Osher und Shen . . . . .	52

3.1.1	Das klassische <i>TV restoration model</i> von Rudin, Osher und Fatemi . .	52
3.1.2	Filterbeschreibung . . . . .	55
3.1.3	Zusammenhang zwischen diskretem Filter und kontinuierlicher TV– Diffusionsgleichung . . . . .	56
3.1.4	Numerische Ergebnisse für die Interpretation im Kontext hyperboli- scher Erhaltungsgleichungen . . . . .	58
3.2	TV–Filter von Marquina und Osher . . . . .	69
3.2.1	Grundlegende Gleichungen und kontinuierliches Filtermodell . . . . .	69
3.2.2	Diskretisierung und numerische Umsetzung . . . . .	71
3.3	Neue Diskretisierungen der Filtermodelle . . . . .	82
3.3.1	Diskretisierungen für die Euler–Lagrange–Gleichung . . . . .	83
3.3.2	Diskretisierungen für die zeitabhängige Modellgleichung . . . . .	86
3.3.3	Diskretisierung der Modellgleichung von Marquina und Osher . . . . .	86
3.3.4	Zeitabhängige Version des TV–Filters von Chan, Osher und Shen . . .	87
3.3.5	Zusammenfassung der Beobachtungen . . . . .	88
3.4	Untersuchung der Gleichung $u_t = (\operatorname{sgn} u_x)_x$ . . . . .	89
3.4.1	Konsistente Kettenregel in $1D$ . . . . .	90
3.4.2	Diskretisierungen und numerische Ergebnisse . . . . .	93
<b>4</b>	<b>Entwicklung neuer Filtermethoden</b>	<b>103</b>
4.1	Extrema–Diffusion . . . . .	103
4.1.1	Motivation und Herleitung des kontinuierlichen Filtermodells . . . . .	104
4.1.2	Äquivalenz der konservativen und nichtkonservativen Modellgleichung . . . . .	106
4.1.3	Direkte Diskretisierung der konservativen Filter- gleichung . . . . .	108
4.1.4	Direkte Diskretisierung der nichtkonservativen Filtergleichung . . . . .	112
4.1.5	Konservative Diskretisierung der konservativen Filtergleichung . . . . .	114
4.1.6	Amplituden– und Phasenverhalten der nichtkonservativen Filtergleichung	136
4.2	Prinzip der Signalzerlegung . . . . .	152
4.2.1	Grundlegende Methode . . . . .	152

4.2.2	Diffusionsanalyse . . . . .	153
4.2.3	Zusammenhang zum Soft Wavelet Shrinkage . . . . .	159
4.2.4	Resultierende Algorithmen und ihre Anwendung . . . . .	161
4.3	Vollständig diskrete Filterroutinen . . . . .	171
4.3.1	Diskrete Filterentwicklung auf Grundlage der Algorithmen von Engquist, Lötstedt und Sjögreen . . . . .	172
4.3.2	Erste Filtervariante . . . . .	177
4.3.3	Zweite Variante: Behebung der Filtermängel . . . . .	178
4.3.4	Zusammenfassung der Ergebnisse . . . . .	184
<b>5</b>	<b>Zusammenfassung und Ausblick</b>	<b>187</b>
	<b>Literaturverzeichnis</b>	<b>191</b>
	<b>Zusammenfassung</b>	<b>195</b>





# Einleitung

Die vorliegende Arbeit beschäftigt sich mit dem numerischen Lösen hyperbolischer Systeme von Erhaltungsgleichungen. Dies sind zeitabhängige Systeme nichtlinearer partieller Differentialgleichungen, die das Verhalten physikalischer Transport- oder Advektionsvorgänge modellieren.

Im Allgemeinen ist es nicht möglich, diese Systeme nichtlinearer Gleichungen exakt zu lösen, so daß die Notwendigkeit entsteht, numerische Methoden zu entwickeln, die approximative Lösungen für die gegebenen Problemstellungen liefern. Diese Situation trifft allgemein auf nichtlineare partielle Differentialgleichungen zu, doch die Klasse der hyperbolischen Erhaltungsgleichungen spielt eine übergeordnete Rolle. Einerseits beruhen viele praktische Probleme aus den Natur- und Ingenieurwissenschaften, vor allem aus dem Bereich der Strömungsmechanik, auf hyperbolischen Systemen von Erhaltungsgleichungen, andererseits ist das Verhalten dieser partiellen Differentialgleichungen vom mathematischen Standpunkt aus besonders interessant. Hyperbolische Erhaltungsgleichungen besitzen die Eigenschaft, daß ihre Lösungen, selbst für stetige Anfangswertbedingungen, Unstetigkeiten aufweisen. Von einer Lösung im klassischen Sinne kann daher nicht mehr gesprochen werden, und die Einführung eines neuen Lösungsbegriffs wird erforderlich.

Die Tatsache, daß hyperbolische Erhaltungsgleichungen Unstetigkeiten in ihren Lösungen erzeugen, wirkt sich auch auf die numerische Approximation ihrer Lösungen aus. Einfache Finite-Differenzen-Verfahren, die für stetige Lösungen sehr gute numerische Resultate erzielen, liefern beim Auftreten von Unstetigkeiten nur sehr schlechte oder gar keine approximativen Lösungen, da sich entweder die Approximationsordnung stark verschlechtert oder Instabilitäten auftreten. Typischerweise erzeugen Verfahren erster Ordnung numerische Lösungen, die in der Nähe der Unstetigkeiten zu stark glättend wirken, während Verfahren zweiter oder höherer Ordnung Oszillationen in der Umgebung der Unstetigkeiten verursachen, die die Stabilität des Verfahrens zerstören können. Ziel ist daher die Konstruktion stabiler numerischer Methoden mit hoher Approximationsordnung, die Unstetigkeiten hochaufgelöst darstellen.

Um diese teilweise gegensätzlichen Anforderungen möglichst gut zu erfüllen, wurden in den letzten Jahren verschiedenste technisch hochentwickelte Differenzenverfahren, wie beispielsweise die TVD- [23], ENO- [25, 62] oder WENO-Methoden [59], konstruiert. Sie beruhen auf der grundlegenden Idee, künstliche nichtlineare Diffusionsterme in Verfahren höherer Ordnung zu integrieren, die Oszillationen in der Umgebung von Unstetigkeiten durch Erhöhung der Dissipation reduzieren. Allerdings führen diese hochentwickelten Methoden oft zu implizit definierten Diffusionstermen, die die mathematische Untersuchung der resultierenden Verfahren sehr aufwendig oder unmöglich machen. Darüber hinaus ist der rechentechnische Aufwand dieser Verfahren sehr hoch.

Im Gegensatz zu diesen ausgefeilten aber rechenzeitintensiven Methoden untersuchen wir in der vorliegenden Arbeit diskrete Filterroutinen der Bildverarbeitung, die wir als nichtlineare Diffusionsterme interpretieren. Durch ihre Integration in einfache Differenzenverfahren höherer Ordnung konstruieren wir neue Methoden für die Approximation hyperbolischer Erhaltungsgleichungen. Die resultierenden Verfahren besitzen explizit gegebene Diffusionsterme und berechnen numerische Lösungen aufgrund der Einfachheit der verwendeten Differenzenverfahren mit geringem Aufwand.

Auf dem ersten Blick erscheint die Anwendung von Filterroutinen der Bildverarbeitung im Kontext numerischer Lösungen von Erhaltungsgleichungen fragwürdig, doch bei genauerer Betrachtung läßt sich eine weitgehende Übereinstimmung der im jeweiligen Anwendungsgebiet verfolgten Ziele feststellen. Die Aufgabe der Bildverarbeitung, ein durch Rauschen und Übertragungsfehler gestörtes Bild wiederherzustellen, ohne dabei Bildstrukturen wie Ecken und Kanten zu zerstören, ist vergleichbar mit der Aufgabe, die Oszillationen numerischer Lösungen von Erhaltungsgleichungen zu entfernen, ohne die Unstetigkeiten unnötig stark zu glätten.

Die Tatsache, daß viele Methoden der Bildverarbeitung auf der Theorie numerischer Lösungen von Erhaltungsgleichungen beruhen, bestätigt die Verwandtschaft beider Bereiche und unterstützt die Vermutung, daß auch der umgekehrte Weg sinnvoll eingeschlagen werden kann.

In der Dissertation von Grahs [21] wird diese Richtung erstmals beschritten, indem anisotrope Diffusionsmodelle, die mit dem Perona–Malik–Modell der Bildverarbeitung verwandt sind, erfolgreich in mehrdimensionale numerische Verfahren für Erhaltungsgleichungen eingebaut werden. Ein eindimensionaler Ansatz, der ebenfalls zu einer Diffusion ähnlich dem Perona–Malik–Modell führt, wurde vor kurzem von Wei [74] veröffentlicht, wobei seine resultierenden Verfahren mit den ENO– und WENO–Methoden vergleichbar sind.

Wir folgen diesem neuen Ansatz und untersuchen weitere nichtlineare Filtermethoden der Bildverarbeitung auf ihre Anwendbarkeit auf numerische Lösungen von Erhaltungsgleichungen.

Dabei orientieren wir uns zunächst an zwei diskreten TV–Filtern der Bildverarbeitung, die wir auf den Kontext numerischer Approximationen von Erhaltungsgleichungen übertragen und als nichtlineare Dissipation in einfache Differenzenverfahren höherer Ordnung integrieren. Numerische Simulationen bestätigen die Leistungsfähigkeit dieser Dissipationen, die typischen Oszillationen einfacher numerischer Verfahren höherer Ordnung zu reduzieren, ohne die Unstetigkeiten der Lösungen unnötig stark zu glätten. Selbst das instabile Zentrale–Differenzen–Verfahren zweiter Ordnung läßt sich durch die nichtlinearen Filterroutinen stabilisieren. Allerdings führen einige der neu konstruierten numerischen Methoden zu Verschiebungsfehlern in den Lösungen, die nicht zufriedenstellend eliminiert werden können. Dennoch zeigen diese ersten Resultate, daß die nichtlinearen Filter der Bildverarbeitung geeignete und einfache Dissipationen für Differenzenverfahren höherer Ordnung darstellen können, so daß wir diese Methoden in der vorliegenden Arbeit genauer untersuchen und verschiedenste neue Lösungsverfahren für hyperbolische Erhaltungsgleichungen entwickeln. Dabei berücksichtigen wir sowohl kontinuierliche Filteransätze, die auf Ideen der Bildverarbeitung basieren, als auch vollständig diskrete Methoden. Für die zweidimensionalen Testfälle implementieren wir die resultierenden Verfahren erstmals auf unstrukturierten Gittern.

## Gliederung der Arbeit

Die Arbeit ist in vier Kapitel unterteilt. Im ersten Kapitel werden die grundlegenden Eigenschaften hyperbolischer Erhaltungsgleichungen zusammengefaßt und die numerischen Methoden für ihre Diskretisierung und Lösung erläutert. Darüber hinaus werden die im Verlauf der Arbeit verwendeten Testfälle vorgestellt.

Das zweite Kapitel ist der allgemeinen Herleitung nichtlinearer Filter gewidmet. Einerseits wird auf den Variationsansatz und auf Ansätze basierend auf partiellen Differentialgleichungen als klassische Methoden der Bildverarbeitung eingegangen, andererseits werden die ersten vollständig diskreten Filterkonstruktionen für Erhaltungsgleichungen dargestellt. Mit Hilfe der kontinuierlichen Gleichung, die für einen diskreten Filter entwickelt wird, kann ein Zusammenhang zu den klassischen Filterherleitungen aufgezeigt werden.

In Kapitel drei stellen wir zwei TV-Filter aus der Bildverarbeitung vor, die wir auf die Situation hyperbolischer Erhaltungsgleichungen anpassen und als künstliche Dissipation in einfache Differenzenverfahren höherer Ordnung integrieren. Die daraus resultierenden Lösungsmethoden werden in ein- und zweidimensionale Testfälle implementiert, die die Leistungsfähigkeit dieser neuen Algorithmen demonstrieren. Da einige dieser neuen Verfahren einen Verschiebungsfehler in der Lösung produzieren, der auf die jeweilige Diskretisierung zurückzuführen ist, entwickeln wir für die zugrundeliegende kontinuierliche Gleichung eine Reihe neuer Diskretisierungen und untersuchen ihre numerischen Eigenschaften. Da die resultierenden Lösungsverfahren stets eine Diskretisierung der Gleichung  $u_t = (\operatorname{sgn} u_x)_x$  enthalten, wird diese Gleichung im letzten Teil des Kapitels gesondert betrachtet, und ihre mathematischen und numerischen Eigenschaften werden analysiert.

Im vierten Kapitel entwickeln wir, aufbauend auf Ideen der Bildverarbeitung und der vollständig diskreten Filterkonstruktion, verschiedene neue Filteralgorithmen für die numerische Approximation hyperbolischer Erhaltungsgleichungen.

Im ersten Abschnitt konstruieren wir eine kontinuierliche Gleichung, die als Extrema-Diffusion interpretiert werden kann, und demonstrieren verschiedene Diskretisierungsmöglichkeiten mit ihren zugehörigen numerischen Resultaten. Wie bereits in Kapitel drei beobachtet, produzieren auch hier einige Diskretisierungen Verschiebungsfehler in der numerischen Lösung. Da die Filteralgorithmen eine sehr einfache Struktur aufweisen, kann eine von Neumann-Analyse durchgeführt werden, die die Entstehung des Verschiebungsfehlers allerdings nicht erklärt.

Im zweiten Teil des Kapitels stellen wir eine neue Implementierungsmöglichkeit der als künstliche Dissipation betrachteten nichtlinearen Filter vor, die auf einer Zerlegung des Eingangsdatensatzes in einen hoch- und einen niederfrequenten Anteil basiert. Wir zeigen, daß das Verfahren qualitativ mit dem Soft Wavelet Shrinkage, einem zurzeit hochaktuellem Forschungsgebiet der Bildverarbeitung, übereinstimmt.

Der letzte Abschnitt ist der Entwicklung neuer vollständig diskreter Filterroutinen gewidmet, die wir in einfache Differenzenverfahren höherer Ordnung integrieren. Die Leistungsfähigkeit der resultierenden Algorithmen wird anhand verschiedener Testfälle für die eindimensionalen Eulergleichungen überprüft.

Eine Zusammenfassung der präsentierten Ergebnisse sowie ein Ausblick auf weiterführende Zielsetzungen schließt die Arbeit ab.



# 1 Grundlagen

Wie in der Einleitung dargestellt, beschäftigen wir uns in der vorliegenden Arbeit mit der Integration diskreter, nichtlinearer Filter der Bildverarbeitung in einfache Differenzenverfahren höherer Ordnung, um auf diese Weise neue Methoden für die numerische Lösung hyperbolischer Erhaltungsgleichungen zu entwickeln.

Bevor wir auf die Konstruktion dieser Algorithmen näher eingehen, geben wir zunächst einen kurzen Überblick über die grundlegenden Eigenschaften hyperbolischer Erhaltungsgleichungen und ihrer numerischen Approximation.

Darüber hinaus stellen wir am Ende des Kapitels die im Verlauf der Arbeit verwendeten Testfälle für die numerischen Simulationen vor.

## 1.1 Hyperbolische Erhaltungsgleichungen

Viele Problemstellungen der Natur- und Ingenieurwissenschaften, insbesondere der Bereich der Gasdynamik, führen zu der Untersuchung nichtlinearer hyperbolischer Erhaltungsgleichungen. Die Tatsache, daß die Lösungen hyperbolischer Gleichungen selbst bei stetigen Anfangs- oder Randwerten unstetig sein können, ist verantwortlich für das große Interesse, daß diesem Gebiet seit Jahrzehnten entgegen gebracht wird. Klassische Eindeutigkeits- und Existenzaussagen für Lösungen hyperbolischer Erhaltungsgleichungen sind aufgrund der entstehenden Unstetigkeiten nicht länger anwendbar und eine neue Definition des Lösungsbegriffs ist notwendig.

In einer Raumdimension sind die hyperbolischen Erhaltungsgleichungen durch

$$\frac{\partial}{\partial t}u(x, t) + \frac{\partial}{\partial x}f(u(x, t)) = 0 \quad (1.1)$$

gegeben, wobei  $u : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}^m$  der  $m$ -dimensionale Vektor der Erhaltungs- oder Zustandsgrößen ist. In einem strömungsdynamischen Problem sind dies beispielsweise Masse, Impuls und Energie. Für die  $i$ te Zustandsvariable bezeichnet  $u_i$  die zugehörige Dichtefunktion, das heißt, daß

$$\int_{x_1}^{x_2} u_i(x, t) dx$$

die Gesamtmenge dieser Zustandsgröße im Intervall  $[x_1, x_2]$  zum Zeitpunkt  $t$  darstellt. Die Erhaltungseigenschaft bedeutet dabei, daß  $\int_{-\infty}^{\infty} u_i(x, t) dx$  hinsichtlich  $t$  konstant ist. Die

Funktion  $u_i$ , die die räumliche Verteilung der Zustandsgröße zur Zeit  $t$  darstellt, ändert sich dagegen im Laufe der Zeit. Die Funktion  $f(u)$ ,  $f : \mathbb{R}^m \rightarrow \mathbb{R}^m$ , heißt *Flußfunktion*. Für einen gegebenen Punkt  $(x, t)$  beschreibt sie die Durchflußmenge, kurz den *Fluß*, jeder Zustandsgröße, wobei der Fluß der  $i$ ten Komponente durch die Funktion  $f_i(u(x, t))$  gegeben ist. Für Gleichung (1.1) ist vorauszusetzen, daß sich der Fluß jeder Zustandsvariablen für  $(x, t)$  berechnen läßt, sobald die Verteilung  $u(x, t)$  bekannt ist. Die Flußfunktionen sind im Allgemeinen nichtlineare Funktionen von  $u$ , was zu nichtlinearen Systemen partieller Differentialgleichungen führt.

Um gegebene Problemstellungen mit Hilfe der Erhaltungsgleichungen zu modellieren, sind zusätzlich zu Gleichung (1.1) bestimmte Anfangs- oder Randwertbedingungen notwendig, die aus der Problemstellung hergeleitet werden müssen. Das einfachste Problem ist das reine Anfangswert- oder *Cauchy-Problem*, das (1.1) für alle  $-\infty < x < \infty$  und  $t \geq 0$  erfüllt und dessen Anfangswert gemäß

$$u(x, 0) = u_0(x), \quad -\infty < x < \infty,$$

gegeben ist.

Die Hyperbolizität von (1.1) ist abhängig von der Jacobi-Matrix  $f'(u)$  der Flußfunktion. Sind die Eigenwerte von  $f'(u)$  für jedes  $u$  reell und existiert eine vollständige Menge von  $m$  linear unabhängigen Eigenvektoren, das heißt, ist die Matrix  $f'(u)$  diagonalisierbar, so ist das System (1.1) hyperbolisch. Im Gegensatz zu parabolischen und elliptischen partiellen Differentialgleichungen, die diffusive beziehungsweise ausgleichende Prozesse beschreiben, modellieren hyperbolische Gleichungen Transport- und Advektionsvorgänge.

Der Name *Erhaltungsgleichungen* beruht darauf, daß diese Gleichungen die Erhaltungsprinzipien der Physik, die Erhaltung der Masse, des Impulses und der Energie in einem abgeschlossenen System, verkörpern. Um zu verdeutlichen, wie die Erhaltungsgleichungen aus den Gesetzen der Physik abgeleitet werden, entwickeln wir im Folgenden die Gleichung für die Massenerhaltung in einem eindimensionalen Flußproblem.

Als Beispiel für ein solches Flußproblem können wir uns das Strömen eines Gases durch ein Rohr vorstellen. Bezeichnet die skalare Zustandsgröße  $\rho(x, t)$  die Gasdichte an der Stelle  $x$  zum Zeitpunkt  $t$ , so ist die Gesamtmasse  $m(t)$  des Gases in jedem beliebigen Intervall  $[x_1, x_2]$  durch

$$m(t) = \int_{x_1}^{x_2} \rho(x, t) dx$$

gegeben. Bei undurchdringlichen Rohrwänden und der Annahme, daß Masse weder neu entstehen noch zerstört werden kann, ist die Massenänderung in  $[x_1, x_2]$  lediglich darauf zurückzuführen, daß Gas über die Intervallränder  $x_1$  oder  $x_2$  strömt. Bezeichnet  $v(x, t)$  die Geschwindigkeit des Gases an der Stelle  $x$  zur Zeit  $t$ , ist der Gasfluß an diesem Punkt durch  $\rho(x, t)v(x, t)$  gegeben, und die Massenänderung im Intervall  $[x_1, x_2]$  wird durch die Differenz der Flüsse in  $x_1$  und  $x_2$  beschrieben, so daß

$$\frac{d}{dt} \int_{x_1}^{x_2} \rho(x, t) dx = \rho(x_1, t)v(x_1, t) - \rho(x_2, t)v(x_2, t) \quad (1.2)$$

gilt. Durch Zeitintegration von  $t_1$  nach  $t_2$  läßt sich diese Gleichung weiter umformen zu

$$\int_{x_1}^{x_2} \rho(x, t_2) dx = \int_{x_1}^{x_2} \rho(x, t_1) dx + \int_{t_1}^{t_2} \rho(x_1, t)v(x_1, t) dt - \int_{t_1}^{t_2} \rho(x_2, t)v(x_2, t) dt \quad (1.3)$$

und ergibt einen Ausdruck für die Masse in  $[x_1, x_2]$  zur Zeit  $t_2 > t_1$  in Abhängigkeit von  $m(t_1)$  und dem Gesamtfluß an den Randpunkten während dieser Zeitspanne. Die Gleichungen (1.2) und (1.3) heißen *Integralformen* der Erhaltungsgleichung. Für die zugehörige *Differentialform* sind die Umformungen

$$\rho(x, t_2) - \rho(x, t_1) = \int_{t_1}^{t_2} \frac{\partial}{\partial t} \rho(x, t) dt$$

und

$$\rho(x_2, t)v(x_2, t) - \rho(x_1, t)v(x_1, t) = \int_{x_1}^{x_2} \frac{\partial}{\partial x} (\rho(x, t)v(x, t)) dx$$

in (1.3) einzusetzen, wobei  $\rho(x, t)$  und  $v(x, t)$  differenzierbar sein müssen. Es ergibt sich

$$\int_{t_1}^{t_2} \int_{x_1}^{x_2} \left\{ \frac{\partial}{\partial t} \rho(x, t) + \frac{\partial}{\partial x} (\rho(x, t)v(x, t)) \right\} dx dt = 0, \quad (1.4)$$

und da dies für beliebige Intervalle  $[x_1, x_2]$  und  $[t_1, t_2]$  gelten muß, folgt, daß der Integrand von (1.4) Null ist, also

$$\frac{\partial}{\partial t} \rho(x, t) + \frac{\partial}{\partial x} (\rho(x, t)v(x, t)) = 0,$$

beziehungsweise in vereinfachter Schreibweise

$$\rho_t + (\rho v)_x = 0. \quad (1.5)$$

Dies ist die *Differentialform* der Erhaltungsgleichung für die Massenerhaltung. Ist die Geschwindigkeit  $v(x, t)$  eine Funktion von  $\rho(x, t)$ , so ist  $\rho v$  nur noch von  $\rho$  abhängig,  $\rho v = f(\rho)$ , und Gleichung (1.5) wird zu der skalaren Erhaltungsgleichung

$$\rho_t + f(\rho)_x = 0.$$

Auf die Herleitung der Erhaltungsgleichungen für Impuls und Energie wird an dieser Stelle verzichtet, eine detaillierte Darstellung findet sich beispielsweise in den Büchern [26] und [39].

Im Folgenden stellen wir kurz die grundlegenden Eigenschaften hyperbolischer Erhaltungsgleichungen dar. Zur Vereinfachung gehen wir dabei zunächst von dem eindimensionalen skalaren Fall aus.

### 1.1.1 Skalare Erhaltungsgleichungen

#### Lineare Transportgleichung

Die einfachste skalare Erhaltungsgleichung ist die *lineare Transport-* oder *Advektionsgleichung*. Sie ergibt sich aus der Massenerhaltungsgleichung (1.5), wenn die Geschwindigkeit  $v$  als konstant angenommen wird. Um sich von der Vorstellung der Dichte zu lösen, führen wir die allgemeineren Bezeichnungen  $u$  und  $a$  ein und erhalten die lineare Transportgleichung

$$u_t + au_x = 0. \quad (1.6)$$

Das zugehörige Cauchy-Problem ist auf  $-\infty < x < \infty$ ,  $t \geq 0$  und durch die Anfangswertbedingung  $u(x, 0) = u_0$  definiert. Die Lösung dieses Problems für  $t \geq 0$  ist

$$u(x, t) = u_0(x - at),$$

das heißt, der Anfangsdatensatz wird abhängig vom Vorzeichen von  $a$  lediglich nach links oder rechts transportiert. Entlang der *Charakteristiken*  $x - at = x_0$  ist die Lösung  $u(x, t)$  konstant.

Auch für einen räumlich variablen Koeffizienten besitzt die Transportgleichung

$$u_t + (a(x)u)_x = 0$$

mit stetiger Funktion  $a(x)$  Charakteristiken, die durch Kurven, die

$$\begin{aligned} x'(t) &= a(x(t)), \\ x(0) &= x_0 \end{aligned}$$

erfüllen, gegeben sind. Allerdings ist die Lösung  $u(x, t)$  entlang dieser Kurven nicht mehr konstant, sondern breitet sich mit veränderlicher Geschwindigkeit  $a(x)$  aus. Dennoch kann  $u(x, t)$  entlang der Charakteristiken durch einfaches Lösen gewöhnlicher Differentialgleichungen berechnet werden. Gilt dabei  $u_0(x) \in C^k(-\infty, \infty)$ , so ist auch die Lösung  $u(x, t)$  glatt, das heißt,  $u \in C^k((-\infty, \infty) \times (0, \infty))$ .

## Nichtlineare Gleichungen

Im Folgenden betrachten wir die nichtlineare skalare Erhaltungsgleichung

$$u_t + f(u)_x = 0, \quad u(x, 0) = u_0(x) \quad (1.7)$$

für die  $f(u)$  eine nichtlineare, konvexe Funktion von  $u$  darstellt.

Das bekannteste Beispiel dieser Klasse von Erhaltungsgleichungen ist *Burgers' Gleichung*, für die  $f(u) = \frac{1}{2}u^2$  gilt, so daß (1.7) die Form

$$u_t + uu_x = 0 \quad (1.8)$$

annimmt. Anhand dieser Gleichung läßt sich zeigen, daß die Existenz einer klassischen Lösung selbst bei glatten Funktionen  $f$  und  $u_0$  nicht mehr für alle  $t$  gewährleistet ist.

Die Charakteristiken erfüllen

$$x'(t) = u(x(t), t)$$

und  $u(x, t)$  ist entlang dieser Kurven konstant. Daher ist auch  $x'(t)$  konstant, so daß die Charakteristiken Geraden sind, die durch die Anfangswerte bestimmt werden. Ist  $u_0$  eine glatte Funktion und  $t$  klein genug, kann die Lösung von (1.8) durch Lösen von

$$x = \xi + u(\xi, 0)t \quad (1.9)$$

nach  $\xi$  und Einsetzen gemäß  $u(x, t) = u(\xi, 0)$  berechnet werden. Für größeres  $t$  können sich die Charakteristiken dagegen schneiden, beispielsweise wenn ein  $x$  mit  $u_x(x, 0) < 0$  existiert, so daß (1.9) nicht länger eindeutig lösbar ist. Zu dem Zeitpunkt, an dem sich die Charakteristiken das erste Mal schneiden, wird die Steigung von  $u(x, t)$  unendlich groß und ein Stoß entsteht. Eine klassische Lösung der partiellen Differentialgleichung existiert ab diesem Zeitpunkt nicht mehr, und die Notwendigkeit der Definition eines neuen Lösungsbegriffs entsteht.



## Schwache Lösungen

Die Konstruktion schwacher Lösungen läuft darauf hinaus, den klassischen Lösungsbegriff soweit zu verallgemeinern, daß Lösungen von (1.7) nicht mehr differenzierbar sein müssen. Dies läßt sich erreichen, indem die Differentialgleichung in eine Integralformulierung umgeformt wird, die weniger glatte Lösungen zuläßt. Dazu wird die Differentialgleichung mit einer glatten Testfunktion multipliziert, anschließend wird die Gleichung integriert und mit Hilfe der partiellen Integration werden die Ableitungen von der Funktion  $u$  auf die Testfunktion übertragen.

Als Testfunktion verwenden wir Funktionen  $\phi \in C_0^1(\mathbb{R} \times \mathbb{R}^+)$ , wobei  $C_0^1$  die Menge stetig differenzierbarer Funktionen mit kompaktem Träger bezeichnet. Die Multiplikation von  $u_t + f(u)_x = 0$  mit  $\phi(x, t)$  und Integration über  $x$  und  $t$  liefert

$$\int_0^\infty \int_{-\infty}^\infty (\phi u_t + \phi f(u)_x) dx dt = 0.$$

Mit partieller Integration ergibt sich daraus

$$\int_0^\infty \int_{-\infty}^\infty (\phi_t u + \phi_x f(u)) dx dt = - \int_{-\infty}^\infty \phi(x, 0) u(x, 0) dx, \quad (1.10)$$

da die meisten Randterme aufgrund des kompakten Trägers von  $\phi$  wegfallen, und es gilt

### Definition 1.1

Die Funktion  $u(x, t)$  heißt schwache Lösung der Erhaltungsgleichung (1.7), wenn sie (1.10) für alle Testfunktionen  $\phi \in C_0^1(\mathbb{R} \times \mathbb{R}^+)$  erfüllt.

### Bemerkung 1.2

Eine klassische Lösung des Cauchy-Problems (1.7) ist auch eine schwache Lösung dieses Problems, und eine schwache Lösung  $u$  von (1.7), die in  $C^1(\mathbb{R} \times \mathbb{R}^+)$  liegt, ist auch eine klassische Lösung von (1.7).

Weiterführende Betrachtungen zu schwachen Lösungen von Erhaltungsgleichungen sind zum Beispiel in [41] zu finden.

Ungünstigerweise sind schwache Lösungen meistens nicht eindeutig bestimmt, so daß sich die Frage stellt, welche schwache Lösung die physikalisch korrekte ist. Darüber hinaus ist nicht jede Art der Unstetigkeit für eine schwache Lösung erlaubt. Bevor wir auf Sprung- und Entropiebedingungen eingehen, die für diese Probleme Abhilfe schaffen, verdeutlichen wir die Problematik schwacher Lösungen kurz anhand des Riemann-Problems.

### Beispiel 1.3 (Riemann-Problem)

Eine Erhaltungsgleichung zusammen mit einem stückweise konstanten Anfangsdatensatz und nur einer Unstetigkeit wird Riemann-Problem genannt.

Ein klassisches Beispiel für ein solches Problem ist Burgers' Gleichung

$$u_t + uu_x = 0$$

mit der stückweise konstanten Anfangswertbedingung

$$u(x, 0) = \begin{cases} u_l & : x < 0 \\ u_r & : x > 0 \end{cases} . \quad (1.11)$$

Die Lösung der partiellen Differentialgleichung ist abhängig von der Beziehung zwischen  $u_l$  und  $u_r$ . Gilt  $u_l > u_r$ , besitzt das Riemann–Problem die eindeutige schwache Lösung

$$u(x, t) = \begin{cases} u_l & : x < st \\ u_r & : x > st \end{cases}, \quad (1.12)$$

wobei  $s$  durch

$$s = \frac{u_l + u_r}{2} \quad (1.13)$$

gegeben ist und Stoßgeschwindigkeit heißt. Dies ist die Geschwindigkeit mit der sich die Unstetigkeit fortbewegt.

Ist dagegen  $u_l < u_r$ , so existieren unendlich viele schwache Lösungen, darunter auch (1.12). Eine weitere schwache Lösung ist die Verdünnungswelle

$$u(x, t) = \begin{cases} u_l & : x < u_l t \\ x/t & : u_l t \leq x \leq u_r t \\ u_r & : x > u_r t \end{cases},$$

die im Gegensatz zu allen anderen möglichen Lösungen die physikalisch richtige darstellt.

## Stoßgeschwindigkeit

Wie leicht gezeigt werden kann, ist (1.12) nur dann eine schwache Lösung für Burgers' Gleichung, wenn die Stoßgeschwindigkeit durch (1.13) gegeben ist, alle anderen Geschwindigkeiten liefern keine zulässige schwache Lösung. Auch für allgemeines  $f(u)$  existiert eine solche Stoßbedingung, die *Rankine–Hugoniot–Bedingung*

$$f(u_l) - f(u_r) = s(u_l - u_r), \quad (1.14)$$

die für skalare Probleme die Form

$$s = \frac{f(u_l) - f(u_r)}{u_l - u_r} = \frac{[f]}{[u]} \quad (1.15)$$

annimmt, wobei  $[\cdot]$  den Sprung der jeweiligen Größe über die Unstetigkeit hinweg bezeichnet. Die Rankine–Hugoniot–Bedingung schränkt damit die Menge der zulässigen Sprungfunktionen ein, eine Lösung ist nur dann eine schwache Lösung, wenn ihre Stoßgeschwindigkeit (1.15) erfüllt. Dies gilt nicht nur für stückweise konstante Riemann–Probleme mit einer Unstetigkeit, sondern für beliebige sich ausbreitenden Stöße, wobei  $u_l$  und  $u_r$  in diesem Fall die Werte direkt links und rechts vom Stoß bezeichnen und  $s$  die momentane Geschwindigkeit ist, die sich mit  $u_l$  und  $u_r$  ändern kann.

## Entropie–Bedingungen

Wie das obige Beispiel zeigt, sind schwache Lösungen partieller Differentialgleichungen nicht unbedingt eindeutig bestimmt, so daß zusätzliche Bedingungen notwendig sind, um aus den möglichen schwachen Lösungen die physikalisch richtige herauszusuchen. Diese Aufgabe

erfüllen die *Entropie-Bedingungen*. Sie werden über die zugehörigen Charakteristiken hergeleitet und beruhen auf der Tatsache, daß die Charakteristiken eines Stoßes bei fortschreitender Zeit in den Stoß laufen. Für skalare Erhaltungsgleichungen ergibt sich daraus

**Entropie-Bedingung I:** *Eine Unstetigkeit, die sich mit der Geschwindigkeit*

$$s = \frac{f(u_l) - f(u_r)}{u_l - u_r}$$

*ausbreitet, erfüllt die Entropie-Bedingung und ist damit die physikalisch relevante Lösung, wenn*

$$f'(u_l) > s > f'(u_r) \quad (1.16)$$

*gilt.*

$f'(u)$  ist dabei die charakteristische Geschwindigkeit, und für konvexes  $f$  reduziert sich diese Bedingung auf die Forderung  $u_l > u_r$ .

Für nichtkonvexe skalare Flußfunktionen  $f$  gilt die von Oleinik [45] entwickelte allgemeinere Bedingung

**Entropie-Bedingung II:**  $u(x, t)$  ist die Entropie-Lösung, wenn alle Unstetigkeiten die Eigenschaft

$$\frac{f(u) - f(u_l)}{u - u_l} \geq s \geq \frac{f(u) - f(u_r)}{u - u_r} \quad (1.17)$$

für alle  $u$  zwischen  $u_l$  und  $u_r$  erfüllen.

Ist  $f$  konvex, reduziert sich diese Formulierung wieder auf Bedingung (1.16).

Weitere Versionen der Entropie-Bedingung und die zugehörigen Beweise sind in [45] und [60] zu finden.

Eine andere Möglichkeit, aus der Menge schwacher Lösungen die physikalisch relevante zu erhalten, bietet die *Entropie-Ungleichung*. Sie beruht auf dem Prinzip der Gasdynamik, daß die physikalische Entropie für Partikel, die im stetigen Fluß sind, konstant ist und ansteigt, wenn eine Unstetigkeit passiert wird. Diese Gesetzmäßigkeit erlaubt die Konstruktion einer Entropiefunktion und eines Entropieflusses, die für glatte Lösungen eine weitere Erhaltungsgleichung erfüllen, die bei unstetigen Lösungen zu einer Ungleichung, der Entropie-Ungleichung wird. Für die mathematische Herleitung der daraus resultierenden Entropie-Bedingung sei auf die Arbeiten von Kruzkov [33] und Lax [34] verwiesen, hier soll lediglich das Resultat formuliert werden:

**Entropie-Bedingung III:**  $u(x, t)$  ist die Entropie-Lösung, wenn für alle konvexen Entropiefunktionen  $\nu(u)$  und zugehörigen Entropieflüsse  $\psi(u)$  die Ungleichung

$$\nu(u)_t + \psi(u)_x \leq 0$$

im schwachen Sinne erfüllt wird.

### 1.1.2 Systeme von Erhaltungsgleichungen

Da die besonderen Eigenschaften hyperbolischer Erhaltungsgleichungen und die daraus folgende Lösungsproblematik schon für den Fall skalarer Gleichungen deutlich geworden ist, verzichten wir an dieser Stelle auf die entsprechenden Darstellungen für Systeme von Erhaltungsgleichungen. Alle im vorausgehenden Kapitel aufgeführten Ergebnisse über skalare Erhaltungsgleichungen und ihre Lösungen lassen sich auf Systeme hyperbolischer Erhaltungseigenschaften verallgemeinern. Ausführliche Darstellungen darüber finden sich in jedem Lehrbuch über hyperbolische Erhaltungseigenschaften, unter anderem in [19, 26, 32, 39, 40].

#### Eulergleichungen der Gasdynamik

Das bedeutendste Beispiel für hyperbolische Systeme nichtlinearer Erhaltungsgleichungen sind die Eulergleichungen der Gasdynamik, auf die wir an dieser Stelle kurz eingehen wollen. Sie beschreiben das Strömungsverhalten eines kompressiblen Gases oder einer kompressiblen Flüssigkeit bei hohem Druck, so daß Reibungseffekte vernachlässigt werden können. Mathematisch interessant sind die Eulergleichungen, da sie, ebenso wie das skalare Cauchy-Problem für Burgers' Gleichung, trotz stetiger Anfangsbedingungen nach endlicher Zeit unstetige Lösungen aufweisen.

Wir haben bereits die Massenerhaltungsgleichung (1.5),

$$\rho_t + (\rho v)_x = 0, \quad (1.18)$$

hergeleitet, die die erste grundlegende Gleichung für die Eulergleichungen darstellt. Die Variable  $\rho$  bezeichnet die Dichte,  $v$  die Geschwindigkeit des betrachteten Gases, das weiterhin durch die Gesamtenergie  $E$  und den Druck  $p$  beschrieben wird. Ähnlich wie (1.18) ergibt sich für eine Raumdimension die Impulserhaltungsgleichung

$$(\rho v)_t + (\rho v^2 + p)_x = 0 \quad (1.19)$$

und die Erhaltungsgleichung für die Energie

$$E_t + [v(E + p)]_x = 0. \quad (1.20)$$

Insgesamt liefern diese drei Erhaltungssätze das *System der Eulergleichungen*

$$\begin{bmatrix} \rho \\ \rho v \\ E \end{bmatrix}_t + \begin{bmatrix} \rho v \\ \rho v^2 + p \\ v(E + p) \end{bmatrix}_x = 0. \quad (1.21)$$

Für zwei Raumdimensionen ist die Geschwindigkeit des Gases oder der Flüssigkeit durch die Komponenten  $(u, v)$  gegeben, so daß die Eulergleichungen in diesem Fall die Form

$$\begin{bmatrix} \rho \\ \rho u \\ \rho v \\ E \end{bmatrix}_t + \begin{bmatrix} \rho u \\ \rho u^2 + p \\ \rho uv \\ u(E + p) \end{bmatrix}_x + \begin{bmatrix} \rho v \\ \rho uv \\ \rho v^2 + p \\ v(E + p) \end{bmatrix}_y = 0 \quad (1.22)$$

annehmen. Da die unbekannten Funktionen  $\rho$ ,  $u$ ,  $v$ ,  $p$  und  $E$  bisher durch die vier Gleichungen in (1.22) beschrieben werden, ist eine weitere Gleichung für das Lösen der Eulergleichungen notwendig. Sie ergibt sich aus der Zustandsgleichung

$$p = f(\rho, u, v, E),$$

die für ideale Gase, die betrachtet werden sollen, durch

$$p = (\gamma - 1) \left[ E - \frac{\rho}{2}(u^2 + v^2) \right] \quad (1.23)$$

gegeben ist, wobei  $\gamma$  die adiabatische Konstante bezeichnet. Die Funktionen  $\rho$ ,  $u$ ,  $v$  und  $p$  heißen *primitive Variablen* und  $\rho$ ,  $\rho u$ ,  $\rho v$  und  $E$  werden *konservative Variablen* genannt. Eine ausführlichere Beschreibung der Eulergleichungen und der zugrundeliegenden physikalischen Eigenschaften sind beispielsweise bei Whitham [78] und Courant und Friedrichs [13] nachzulesen.

## 1.2 Numerische Methoden

Im Folgenden geben wir einen Überblick über die grundlegenden Konzepte der numerischen Approximation hyperbolischer Erhaltungsgleichungen mittels Finiten-Differenzen- beziehungsweise Finiten-Volumen-Verfahren.

Wird die numerische Berechnung unstetiger Lösungen für Erhaltungsgleichungen unter Verwendung standardmäßiger Diskretisierungen, die für glatte Lösungen entwickelt wurden, vorgenommen, ergeben sich typischerweise sehr schlechte numerische Ergebnisse. Verfahren erster Ordnung liefern Lösungen, die im Bereich der Unstetigkeiten viel zu stark geglättet sind, Verfahren zweiter Ordnung leiden an dispersiven Effekten, die in der numerischen Lösung zu starken Oszillationen in der Umgebung der Unstetigkeiten führen. Gesucht wird daher nach numerischen Methoden, die die Vorteile der Verfahren erster und zweiter Ordnung miteinander vereinen, ohne ihre jeweiligen Nachteile zu übernehmen. Ziel ist es dabei, in glatten Regionen unstetiger Lösungen Ergebnisse von zweiter Ordnung Genauigkeit zu erhalten, Unstetigkeiten scharf aufzulösen und sowohl zu starke Glättung als auch Oszillationen in der Umgebung der Unstetigkeiten zu vermeiden. Da sich diese Eigenschaften teilweise widersprechen, ist es die Aufgabe, numerische Verfahren zu konstruieren, die diese Anforderungen bestmöglichst erfüllen.

Der Großteil der in diesem Kontext entwickelten numerischen Methoden basiert auf eindimensionalen Verfahren, die sich für Systeme von Erhaltungsgleichungen verallgemeinern lassen, indem sie komponentenweise angewendet werden.

Aufgrund dieser Dimensionszerlegung betrachten wir in dem folgenden Abschnitt die wesentlichen Approximationstechniken lediglich für skalare Erhaltungsgleichungen in einer Raumdimension.

### 1.2.1 Grundlegende Diskretisierungen und Eigenschaften

Wir betrachten für eindimensionale skalare Erhaltungsgleichungen das Cauchy–Problem

$$\begin{aligned} u_t + f(u)_x &= 0 \text{ auf } \mathbb{R} \times \mathbb{R}^+, \\ u(x, 0) &= u_0(x) \text{ auf } \mathbb{R}. \end{aligned} \quad (1.24)$$

Die  $x$ – $t$ –Ebene diskretisieren wir durch die Wahl einer *Gitterweite*  $h \equiv \Delta x$  und einer *Zeitschrittweite*  $k \equiv \Delta t$ , wobei wir zur Vereinfachung von einem gleichmäßigen Gitter ausgehen, so daß  $h$  und  $k$  konstant sind. Die diskreten Gitterpunkte  $(x_i, t_n)$  werden durch

$$\begin{aligned} x_i &= ih, \quad i = \dots, -1, 0, 1, \dots, \\ t_n &= nk, \quad n = 0, 1, 2, \dots, \end{aligned}$$

definiert, und die tatsächliche Lösung  $u$  an diesen Punkten wird mit

$$u_i^n = u(x_i, t_n)$$

bezeichnet. Die numerische Lösung  $U_i^n$  kann auf klassische Weise als Approximation von  $u_i^n$  am Gitterpunkt  $(x_i, t_n)$  aufgefaßt werden. Allerdings ist es im Kontext von Erhaltungsgleichungen vorzuziehen,  $U_i^n$  als Approximation des Mittelwerts von  $u(x, t_n)$  auf der zugehörigen Gitterzelle zu interpretieren, wobei dieser Zellmittelwert durch

$$\bar{u}_i^n := \frac{1}{h} \int_{x_{i-1/2}}^{x_{i+1/2}} u(x, t_n) dx \quad (1.25)$$

mit  $x_{i+1/2} := x_i + h/2 = (i + \frac{1}{2})h$  definiert ist. Da die Integralform der Erhaltungsgleichung (1.24) für  $Z_i = [x_{i-1/2}, x_{i+1/2}]$  und das Zeitintervall  $[t_n, t_{n+1}]$  durch

$$\int_{x_{i-1/2}}^{x_{i+1/2}} [u(x, t_{n+1}) - u(x, t_n)] dx + \int_{t_n}^{t_{n+1}} [f(u(x_{i+1/2}, t)) - f(u(x_{i-1/2}, t))] dt = 0 \quad (1.26)$$

gegeben ist und gerade die zeitliche Entwicklung solcher Integrale, wie sie in (1.25) auftreten, beschreibt, ist letztere Interpretation von  $U_i^n$  die natürlichere und wird im Folgenden beibehalten. Analog definieren wir den numerischen Anfangsdatensatz  $U_i^0 := \bar{u}_i^0$ . Die Integralform (1.26) verdeutlicht darüber hinaus die physikalische Erhaltungseigenschaft, daß die Änderung von  $u$  in der Zelle  $Z_i$  im Zeitintervall  $[t_n, t_{n+1}]$  bei der Abwesenheit von Quellen und Senken lediglich von der Differenz der Flüsse durch die Zellränder  $x_{i+1/2}$  und  $x_{i-1/2}$  abhängig ist. Definieren wir die numerischen Approximationen für die Flußintegrale gemäß

$$\begin{aligned} F_{i+1/2}^n &:= F(U_{i-r+1}^n, \dots, U_{i+r}^n) \approx \frac{1}{k} \int_{t_n}^{t_{n+1}} f(u(x_{i+1/2}, t)) dt, \\ F_{i-1/2}^n &:= F(U_{i-r}^n, \dots, U_{i+r-1}^n) \approx \frac{1}{k} \int_{t_n}^{t_{n+1}} f(u(x_{i-1/2}, t)) dt, \end{aligned} \quad (1.27)$$

läßt sich (1.26) mit Hilfe von (1.25) durch

$$\frac{1}{k}(U_i^{n+1} - U_i^n) = \frac{1}{h}(F_{i+1/2}^n - F_{i-1/2}^n) \quad (1.28)$$

approximieren. Dies ist die diskrete Version von (1.26), die sich mit der Definition der *Gitterkonstanten*  $\lambda := k/h$  in das explizite numerische Verfahren

$$U_i^{n+1} = U_i^n - \lambda(F_{i+1/2}^n - F_{i-1/2}^n) \quad (1.29)$$

umformulieren läßt. Zusammen mit den Anfangswerten  $U^0$  ergibt die Differenzengleichung (1.29) eine explizite Methode zur Berechnung der neuen Werte  $U^{n+1}$  aus den zur Zeit  $t^n$  bekannten Zellmittelwerten  $U^n$ .

Darüber hinaus existieren implizite numerische Verfahren, die zur Berechnung von  $U^{n+1}$  sowohl Werte von  $U^n$  als auch von  $U^{n+1}$  verwenden. Da in der vorliegenden Arbeit lediglich explizite Verfahren behandelt werden, betrachten wir die impliziten Methoden im Folgenden nicht weiter.

Für die Bestimmung von  $F_{i+1/2}$  und  $F_{i-1/2}$  in (1.29) ist eine numerische Approximation der Flußfunktion  $f(u)$  notwendig. Als einfachste Möglichkeit bietet sich die Wahl einer zentralen Approximation, die  $F_{i+1/2} = \frac{1}{2}(f(U_{i+1}) + f(U_i))$  ergibt. Das resultierende Verfahren approximiert (1.24) zwar mit der Genauigkeit zweiter Ordnung, ist aber numerisch instabil. Es stellt sich daher die Aufgabe, eine geeignete numerische Approximationen für (1.27) zu konstruieren. Dabei ist nicht nur auf die Approximationsordnung und Stabilität der resultierenden Differenzengleichung zu achten, sondern es sollte auch gewährleistet werden, daß die numerische Lösung gegen die physikalisch korrekte schwache Lösung der Erhaltungsgleichung konvergiert.

### Konservativität eines Differenzenverfahrens

Eine notwendige Bedingung, die die Konvergenz des numerischen Verfahrens gegen eine schwache Lösung, die die Entropie-Bedingung verletzt, verhindert, ist die Darstellung des Verfahrens in *Erhaltungsform*:

#### Definition 1.4

Ein numerisches Verfahren heißt in Erhaltungsform darstellbar, wenn es in der Form

$$\begin{aligned} U_i^{n+1} &= U_i^n - \lambda [F(U_{i-p}^n, U_{i-p+1}^n, \dots, U_{i+q}^n) - F(U_{i-p-1}^n, U_{i-p}^n, \dots, U_{i+q-1}^n)] \\ &= U_i^n - \lambda [F_{i+1/2}^n - F_{i-1/2}^n] \end{aligned} \quad (1.30)$$

angegeben werden kann, wobei  $F$  eine stetige Funktion mit  $p + q + 1$  Argumenten ist.  $F$  heißt numerische Flußfunktion und  $F_{i+1/2}^n = F(U_{i-p}^n, U_{i-p+1}^n, \dots, U_{i+q}^n)$  bezeichnet den Fluß zwischen den Zellen  $Z_i$  und  $Z_{i+1}$  über die Zellbegrenzung  $x_{i+1/2}$ .

Im einfachsten Fall gilt  $p = 0$  und  $q = 1$  und (1.30) ergibt

$$U_i^{n+1} = U_i^n - \lambda [F(U_i^n, U_{i+1}^n) - F(U_{i-1}^n, U_i^n)]. \quad (1.31)$$

#### Beispiel 1.5 (Lax–Friedrichs)

Das Lax–Friedrichs–Verfahren für (1.24) ist durch die Differenzengleichung

$$U_i^{n+1} = \frac{1}{2}(U_{i-1}^n + U_{i+1}^n) - \frac{\lambda}{2} (f(U_{i+1}^n) - f(U_{i-1}^n))$$

gegeben. Mit Hilfe der numerischen Flußfunktion

$$F(U_i, U_{i+1}) = \frac{1}{2\lambda}(U_i - U_{i+1}) + \frac{1}{2}((f(U_i) + f(U_{i+1})))$$

läßt es sich in der konservativen Form (1.31) darstellen.

Verfahren in Erhaltungssform erfüllen darüber hinaus eine globale Konservativitätseigenschaft. Für Indizes  $J < K$ , die hinreichend weit auseinander liegen, und so kleinem  $n$ , daß die numerische Lösung jeweils in einer Umgebung von  $x_{J-1/2}$  und  $x_{K+1/2}$  konstant ist, gilt

$$\int_{x_{J-1/2}}^{x_{K+1/2}} U_k(x, t_n) dx = \int_{x_{J-1/2}}^{x_{K+1/2}} u(x, t_n) dx, \quad (1.32)$$

wobei  $U_k(x, t)$  gemäß  $U_k(x, t) = U_i^n$  für  $(x, t) \in [x_{i-1/2}, x_{i+1/2}) \times [t_n, t_{n+1})$  definiert ist und stückweise konstant ist. Diskrete Verfahren, die eine Darstellung in Erhaltungssform besitzen und damit (1.32) erfüllen, werden *konservativ* genannt.

Die genaue Herleitung von (1.32) aus der Integralform der Erhaltungsgleichung ist beispielsweise in [39] zu finden.

## Konsistenz

Eine weitere, elementare Eigenschaft eines numerischen Verfahrens ist seine Konsistenz mit der zugehörigen Erhaltungsgleichung, also die Approximation der richtigen, zugrundeliegenden Differentialgleichung.

### Definition 1.6

Das numerische Verfahren (1.30) heißt konsistent mit (1.24), wenn für alle  $u \in \mathbb{R}$

$$F(u, \dots, u) = f(u)$$

gilt.

## Lax–Wendroff–Theorem

Für konservative und konsistente Differenzenverfahren garantiert das bekannte Lax–Wendroff–Theorem [36], unter der Voraussetzung, daß das Verfahren bei Gitterverfeinerung mit  $k_l, h_l \rightarrow 0$  gegen eine Funktion  $u(x, t)$  konvergiert, die Konvergenz gegen eine schwache Lösung der zugehörigen Erhaltungsgleichung. Wir geben hier eine von der ursprünglichen Formulierung leicht abweichende aber äquivalente Version des Lax–Wendroff–Theorems an [39]:

### Theorem 1.7 (Lax–Wendroff–Theorem)

Gegeben sei eine Folge von Gittern mit Index  $l = 1, 2, \dots$  und Gitterkonstanten  $k_l, h_l \rightarrow 0$  für  $l \rightarrow \infty$ , wobei  $k_l = \lambda h_l$  für konstantes  $\lambda$  gelte. Die auf dem  $l$ ten Gitter mit einem konservativen und konsistenten Verfahren berechnete numerische Lösung werde mit  $U_l(x, t)$  bezeichnet. Für  $l \rightarrow \infty$  konvergiere  $U_l$  gegen eine Funktion  $u$  im folgenden Sinn:



1. Für jede Menge  $\Omega = [a, b] \times [0, T]$  gelte

$$\int_0^T \int_a^b |U_l(x, t) - u(x, t)| dx dt \rightarrow 0 \text{ für } l \rightarrow \infty$$

oder äquivalent

$$\|U_l - u\|_{1, \Omega} \rightarrow 0 \text{ für } l \rightarrow \infty.$$

2. Zusätzlich existiere für jedes  $T$  ein  $R > 0$ , so daß

$$TV(U_l(\cdot, t)) < R$$

für alle  $0 \leq t \leq T$ ,  $l = 1, 2, \dots$  gelte, wobei

$$TV(v) = \sup \sum_{j=1}^N |v(\xi_j) - v(\xi_{j-1})|$$

die Totalvariation von  $v$  bezeichne, und das Supremum über alle Intervallzerlegungen von  $-\infty = \xi_0 < \xi_1 < \dots < \xi_N = \infty$  zu bilden ist.

Dann ist  $u(x, t)$  eine schwache Lösung der zugehörigen Erhaltungsgleichung.

Das Lax–Wendroff–Theorem gewährleistet lediglich die Konvergenz gegen eine schwache Lösung der Erhaltungsgleichung, nicht die Konvergenz gegen die physikalisch korrekte Entropie–Lösung. Dennoch hat das Theorem für die numerische Simulation große Bedeutung, es besagt, daß eine auf einem festen Gitter berechnete Lösung, die sinnvoll erscheint und Unstetigkeiten gut auflöst, in der Regel auch tatsächlich eine gute Approximation an eine schwache Lösung darstellt.

Da das Theorem nicht garantiert, daß das numerische Verfahren konvergiert, sind weitere Eigenschaften wie Stabilität zu erfüllen, um die Konvergenz zu gewährleisten.

### CFL–Bedingung

Wie bereits festgestellt, breiten sich Lösungen hyperbolischer Erhaltungsgleichungen entlang ihrer Charakteristiken aus, so daß sich durch Rückverfolgung ihrer Charakteristiken die Anfangswerte bestimmen lassen, von denen die Lösung an einer gegebenen Stelle abhängt. Dieser Bereich, der die Lösung beeinflußt, heißt *Abhängigkeitsbereich* der partiellen Differentialgleichung. Entsprechend heißt die Punktmenge, deren Anfangswerte die numerische Lösung an der Stelle  $(x, t)$  beeinflussen, *numerischer Abhängigkeitsbereich*.

Eine notwendige Bedingung für die Stabilität eines numerischen Verfahrens ist die Eigenschaft, daß der numerische Abhängigkeitsbereich des Differenzenverfahrens den Abhängigkeitsbereich der Differentialgleichung einschließt.

Für zentrale Drei–Punkte–Verfahren, bei denen der Wert  $U_i^n$  von den Werten an den Punkten  $x_{i-1}$ ,  $x_i$ ,  $x_{i+1}$  zur Zeit  $t_{n-1}$  abhängt, ergibt diese Anforderung die *Courant–Friedrichs–Lewy (CFL)–Bedingung*

$$\frac{\Delta t}{\Delta x} \max_{u \in U} |f'(u)| = \lambda \max_{u \in U} |f'(u)| \leq 1, \quad (1.33)$$

die eine notwendige Bedingung für die Stabilität dieser Verfahren darstellt. Die Größe auf der linken Seite von (1.33) heißt *Courant-Zahl* oder *CFL-Zahl*. Weitere Details zur CFL-Bedingung sind in der grundlegenden Arbeit von Courant, Friedrichs und Lewy [14] zu finden.

## Diskrete Entropie-Bedingung

Das Lax-Wendroff-Theorem liefert lediglich die Konvergenz gegen eine schwache Lösung der partiellen Differentialgleichung, nicht unbedingt die Konvergenz gegen die physikalisch korrekte Entropie-Lösung. Für die schwachen Lösungen hyperbolischer Erhaltungsgleichungen existieren die in Kapitel 1.1 vorgestellten Entropie-Bedingungen, mit denen sich aus der Menge der schwachen Lösungen die Entropie-Lösung ermitteln läßt. Ein entsprechendes Kriterium ist auch für die numerischen Lösungen vorhanden.

Ist  $u(x, t)$  Entropie-Lösung, so ist die Ungleichung

$$\frac{\partial}{\partial t} \eta(u(x, t)) + \frac{\partial}{\partial x} \psi(u(x, t)) \leq 0 \quad (1.34)$$

für eine skalare Entropie-Funktion  $\eta(u)$  und Entropie-Fluß  $\psi(u)$  im schwachen Sinne zu erfüllen. Äquivalent dazu ist die Gültigkeit von

$$\int_0^\infty \int_{-\infty}^\infty \phi_t \eta(u) + \phi_x \psi(u) \, dx \, dt \leq - \int_{-\infty}^\infty \phi(x, 0) \eta(u(x, 0)) \, dx \quad (1.35)$$

für alle  $\phi \in C_0^1$  mit  $\phi(x, t) \geq 0$  für alle  $x, t$ .

Um zu zeigen, daß eine schwache Lösung  $u(x, t)$ , die sich bei der Gitterverfeinerung als Grenzfunktion von  $U_l(x, t)$  ergibt, die Ungleichung (1.35) erfüllt, reicht es, die Gültigkeit der diskreten Version von (1.34),

$$\eta(U_i^{n+1}) \leq \eta(U_i^n) - \frac{k}{h} [\Psi(U^n; i) - \Psi(U^n; i-1)], \quad (1.36)$$

nachzuprüfen. Die Funktion  $\Psi$  bezeichnet dabei einen numerischen Entropie-Fluß, der gemäß der Definition 1.6 konsistent mit  $\psi$  sein muß, und  $\Psi(U^n; i)$  ist als Approximation des gemittelten Entropie-Flusses

$$\frac{1}{k} \int_{t_n}^{t_{n+1}} \psi(u(x_{i+1/2}, t)) \, dt$$

aufzufassen.

Läßt sich für geeignetes  $\Psi$  zeigen, daß (1.36) gilt, so erfüllt die schwache Lösung, die durch Gitterverfeinerung als Grenzfunktion entsteht, die Entropie-Ungleichung (1.35) und ist damit die physikalisch korrekte Entropie-Lösung.

## Modifizierte Gleichungen

Ein häufig verwendetes Hilfsmittel zur Analyse von Differenzenverfahren und deren Lösungen ist die Approximation der Differenzengleichung durch eine Differentialgleichung, von der das

qualitative Verhalten der numerischen Methode abgeleitet werden kann. Das Vorgehen dazu wird im Folgenden anhand des Lax–Friedrichs–Verfahrens für das lineare Cauchy–Problem

$$\begin{aligned} u_t + Au_x &= 0, \quad -\infty < x < \infty, \quad t \geq 0, \\ u(x, 0) &= u_0(x) \end{aligned} \quad (1.37)$$

erklärt.

Zunächst ist das Differenzenverfahren so darzustellen, daß es als direkte Diskretisierung der Differentialgleichung (1.37) aufzufassen ist. Für unser Beispiel ist das die Form

$$\frac{1}{k} \left[ U_i^{n+1} - \frac{1}{2}(U_{i-1}^n + U_{i+1}^n) \right] + \frac{1}{2h} A[U_{i+1}^n - U_{i-1}^n] = 0.$$

Nun wird jedes  $U_i^n$  durch die zugehörige exakte Lösung ersetzt. Dabei ergibt sich auf der rechten Seite der Gleichung nicht mehr Null sondern der *lokale Abschneidefehler*

$$L_k(x, t) = \frac{1}{k} \left[ u(x, t+k) - \frac{1}{2}(u(x-h, t) + u(x+h, t)) \right] + \frac{1}{2h} A[u(x+h, t) - u(x-h, t)].$$

Unter der Annahme glatter Lösungen läßt sich jeder auftretende Wert von  $u$  in eine Taylorreihe um  $u(x, t)$  entwickeln, und Umsortierung der Terme liefert

$$L_k(x, t) = u_t + Au_x + \frac{1}{2} \left( ku_{tt} - \frac{h^2}{k} u_{xx} \right) + O(h^2), \quad (1.38)$$

wobei zur Vereinfachung  $u \equiv u(x, t)$  gesetzt wurde. Da  $u(x, t)$  die exakte Lösung von  $u_t + Au_x = 0$  darstellt, fallen die ersten beiden Summanden in (1.38) weg. Zusammen mit  $u_{tt} = A^2 u_{xx}$ , was sich mit Hilfe von  $u_t = -Au_x$  ergibt, und der Bezeichnung  $I$  für die  $m \times m$  Einheitsmatrix wird (1.38) zu

$$\begin{aligned} L_k(x, t) &= \frac{1}{2}k \left( A^2 - \frac{h^2}{k^2} I \right) u_{xx}(x, t) + O(k^2) \\ &= O(k). \end{aligned} \quad (1.39)$$

Da der lokale Abschneidefehler damit linear von  $k$  abhängt, ist das Lax–Friedrichs–Verfahren von erster Ordnung genau.

Betrachten wir dagegen  $u(x, t)$  als Lösung der Differentialgleichung

$$u_t + Au_x + \frac{1}{2} \left( ku_{tt} - \frac{h^2}{k} u_{xx} \right) = 0, \quad (1.40)$$

gilt  $L_k(x, t) = O(k^2)$ , und das Lax–Friedrichs–Verfahren liefert eine Approximation zweiter Ordnung für die Lösung von (1.40). Diese Gleichung heißt die *modifizierte Gleichung* für das Lax–Friedrichs–Verfahren.

Für eine einfachere Analyse der modifizierten Gleichung ist es vorteilhaft, den Term  $u_{tt}$  in (1.40) durch Ableitungen nach  $x$  zu ersetzen. Da  $u$  die Differentialgleichung  $u_t = -Au_x$  nicht mehr erfüllt, ist  $u_{tt}$  daher über Ableitungen von (1.40) zu berechnen. Es ergibt sich  $u_{tt} = A^2 u_{xx} + O(k)$  und damit

$$u_t + Au_x = \frac{h^2}{2k} \left( I - \frac{k^2}{h^2} A^2 \right) u_{xx}, \quad (1.41)$$

so daß das Lax–Friedrichs–Verfahren auch für diese modifizierte Gleichung von zweiter Ordnung genau ist.

Die modifizierte Gleichung (1.41) ist als Advektions–Diffusions–Gleichung

$$u_t + Au_x = Du_{xx} \quad (1.42)$$

mit der Diffusionsmatrix

$$D = \frac{h^2}{2k} \left( I - \left( \frac{k}{h} A \right)^2 \right) \quad (1.43)$$

darstellbar. Gleichungen von diesem Typ sind dafür bekannt, daß die Strukturen in ihren Lösungen mit fortschreitender Zeit verwischt werden. Dies erklärt genau die klassische Eigenschaft des Lax–Friedrichs–Verfahrens, Unstetigkeiten in der numerischen Lösung stark zu überglätten. Alle Verfahren erster Ordnung leiden an diesem Phänomen.

Darüber hinaus läßt sich (1.41) beziehungsweise (1.42) mit (1.43) auch wie folgt interpretieren. Beide Darstellungen zeigen, daß der Term niedrigster Ordnung im Abschneidefehler, bezogen auf die ursprüngliche Differentialgleichung  $u_t + Au_x = 0$ , den Faktor  $u_{xx}$  enthält. Dieser Term gleicht dem Viskositätsterm der eindimensionalen Gleichungen für Flüssigkeitsströmungen und wird daher als *künstliche Viskosität* des Lax–Friedrichs–Verfahrens bezeichnet. Um von zusätzlich und absichtlich in Differenzenverfahren integrierter *expliziter künstlicher Viskosität* zu unterscheiden, wird die Viskosität, die naturgemäß im Differenzenverfahren enthalten ist, auch *implizite künstliche Viskosität* genannt. Die künstliche Viskosität bewirkt eine Verringerung aller Gradienten in der Lösung, was zur Glättung aller Strukturen, unabhängig davon, ob sie physikalisch korrekt oder numerisch bedingt sind, führt. Dieser Effekt, der eine direkte Folge gerader Ableitungsterme im Abschneidefehler ist, wird *Dissipation* genannt.

Ein weiterer unphysikalischer Effekt numerischer Verfahren ist die *Dispersion*, die durch ungerade Ableitungsterme im Abschneidefehler erzeugt wird. Differenzenverfahren zweiter Ordnung, wie beispielsweise das Lax–Wendroff–Verfahren, dessen modifizierte Gleichung für (1.37) durch

$$u_t + Au_x = \frac{h^2}{6} A \left( \frac{k^2}{h^2} A^2 - I \right) u_{xxx}$$

gegeben ist, führen zu Dispersionsfehlern in der numerischen Lösung. Ihre modifizierten Gleichungen lassen sich typischer Weise in der Form

$$u_t + Au_x = \mu u_{xxx}, \quad (1.44)$$

also als *dispersive Gleichung*, darstellen. Als Folge der Dispersion werden die Phasenbeziehungen verschiedener Wellen verzerrt, was zu Oszillationen in der numerischen Lösung führt, wie sie für Differenzenverfahren zweiter Ordnung bekannt sind.

Weiterführende Literatur zur Herleitung und Verwendung modifizierter Gleichungen ist bei Warming und Hyett [73] zu finden.

### 1.2.2 Verschiedene Klassen numerischer Verfahren

Es existiert eine Vielzahl verschiedenster Differenzenverfahren für die numerische Approximation hyperbolischer Erhaltungsgleichungen. Ein Großteil der Methoden wird entwickelt,

indem die Ableitungen der Differentialgleichung durch geeignete finite Differenzen approximiert werden. Dabei ist darauf zu achten, daß das resultierende numerische Verfahren die oben genannten Eigenschaften wie Konservativität, Konsistenz und Stabilität weitgehend erfüllt.

Die Differenzenverfahren sind generell in zwei Klassen zu unterteilen, in Verfahren erster Ordnung, deren implizite Viskosität so groß ist, daß die numerischen Lösungen stark überglättet sind, und in Verfahren zweiter oder höherer Ordnung, die dispersive Effekte hervorrufen, die sich durch Oszillationen in der Lösung bemerkbar machen.

### Monotone und TV-vermindernde Verfahren

Ein klassisches Beispiel für Verfahren erster Ordnung sind die monotonen Verfahren, die wie folgt definiert sind:

#### Definition 1.8

Das Differenzenverfahren  $U_i^{n+1} = \mathcal{H}(U^n; i)$  heißt monoton, wenn

$$\frac{\partial}{\partial U_j^n} \mathcal{H}(U^n; i) \geq 0 \quad \forall i, j, U^n$$

gilt.

Dabei bezeichnet  $\mathcal{H}(U^n; i)$  den diskreten Lösungsoperator für explizite zweistufige Verfahren, der durch die Schreibweise  $U^{n+1} = \mathcal{H}_k(U^n)$  beziehungsweise  $U_i^{n+1} = \mathcal{H}_k(U^n; i)$  gegeben ist.

Die Definition besagt, daß die Vergrößerung eines beliebigen  $U_j^n$  nicht zur Folge haben kann, daß der Wert von  $U_i^{n+1}$  abnimmt.

Für monotone Verfahren gilt

#### Theorem 1.9

Monotone Verfahren sind höchstens von erster Ordnung genau.

Um bei der numerischen Approximation unstetiger Lösungen zu verhindern, daß Oszillationen in der Umgebung der Unstetigkeiten auftreten, wird von Differenzenverfahren gefordert, daß sie Monotonie-erhaltend sind.

#### Definition 1.10

Ein numerisches Verfahren heißt Monotonie-erhaltend, wenn aus der Monotonie von  $U^n$  die Monotonie von  $U^{n+1}$  folgt.

Monotone Verfahren erfüllen zwar die Monotonie-Erhaltung, sind aber für die numerische Approximation eher ungeeignet, da sie lediglich von erster Ordnung genau sind und darüber hinaus Unstetigkeit unverhältnismäßig stark glätten.

Um Methoden höherer Ordnung zu konstruieren, ist eine schwächere Stabilitätsbedingung als die Monotonie notwendig. Vor diesem Hintergrund entwickelte Harten [23] die *Totalvariations-vermindernden* Verfahren, die gewährleisten, daß die Totalvariation

$$TV(U^n) = \sum_{i \in \mathbb{Z}} |U_{i+1}^n - U_i^n|$$

mit der Zeit  $t$  nicht ansteigt.

**Definition 1.11**

Ein numerisches Verfahren  $U_i^{n+1} = \mathcal{H}(U^n; i)$  heißt Totalvariations-vermindernd (TVD), wenn

$$TV(U^{n+1}) \leq TV(U^n)$$

für alle Funktionen  $U^n$  gilt.

Dabei ist zu beachten, daß die Bezeichnung Totalvariations-vermindernd ungenau ist, da diese Verfahren die Totalvariation gemäß ihrer Definition mit fortschreitender Zeit erhalten und nicht vermindern.

Der Zusammenhang zwischen monotonen, Monotonie-erhaltenden und Totalvariations-vermindernden Verfahren wird durch folgendes Lemma beschrieben.

**Lemma 1.12**

- (i) *Monotone Verfahren sind Totalvariations-vermindernd.*
- (ii) *TVD-Verfahren sind Monotonie-erhaltend.*

Zwar ist jedes lineare, Monotonie-erhaltende Verfahren, also auch jedes lineare TVD-Verfahren, nur von erster Ordnung genau (vergleiche Theorem 1.13), aber es besteht die Möglichkeit, nichtlineare TVD-Verfahren zu entwickeln, die von höherer Ordnung sind. Die grundlegenden Ideen für ihre Konstruktion werden im Folgenden kurz skizziert.

## Hochauflösende Methoden

Als hochauflösende Methoden werden numerische Verfahren bezeichnet, die für glatte Lösungen mindestens von zweiter Ordnung genau sind und Unstetigkeiten gut auflösen, ohne Oszillationen zu erzeugen. Sie basieren auf einem Verfahren höherer Ordnung, das so modifiziert wird, daß seine numerische Dissipation in der Umgebung von Unstetigkeiten ansteigt. Dies läßt sich erreichen, indem eine explizite künstliche Viskosität, beispielsweise ein zu  $u_{xx}$  proportionaler Diffusionsterm, zu dem Verfahren höherer Ordnung addiert wird. Dabei ist zu beachten, daß der zusätzliche Summand mit  $h, k \rightarrow 0$  schnell genug gegen Null konvergiert, so daß ein konsistentes Verfahren entsteht, das für glatte Lösungsbereiche die ursprüngliche Anpassungsordnung erhält.

Für die skalare Advektionsgleichung  $u_t + au_x = 0$  kann beispielsweise das bekannte Lax-Wendroff-Verfahren gemäß

$$U_i^{n+1} = U_i^n - \frac{\nu}{2}(U_{i+1}^n - U_{i-1}^n) + \frac{1}{2}\nu^2(U_{i+1}^n - 2U_i^n + U_{i-1}^n) + kQ(U_{i+1}^n - 2U_i^n + U_{i-1}^n) \quad (1.45)$$

erweitert werden, wobei  $\nu = \frac{ak}{h}$  gilt und  $Q$  die neu eingeführte künstliche Viskosität bezeichnet. Für hinreichend großes  $Q$  besteht die Hoffnung, den dispersiven Term des Lax-Wendroff-Verfahrens durch die zusätzlich integrierte Dissipation auszugleichen, so daß Oszillationen in der Lösung verhindert werden und das Verfahren Monotonie-erhaltend wird. Allerdings ist (1.45) für konstantes  $Q$  eine lineare Methode zweiter Ordnung, die nicht Monotonie-erhaltend ist, denn es gilt:

**Theorem 1.13 (Godunov)**

*Ein lineares, Monotonie-erhaltendes numerisches Verfahren ist höchstens von erster Ordnung genau.*

Um die an (1.45) gestellten Anforderungen erfüllen zu können, muß der künstliche Viskositätsterm  $Q$  von der numerischen Lösung  $U^n$  abhängen. In der Nähe der Unstetigkeiten ist ein betragsmäßig größeres  $Q$  zur Erhaltung der Monotonie notwendig, als in stetigen Regionen der Lösung, für die die Dissipation möglichst klein gehalten werden soll, damit die Approximationsgenauigkeit nicht verschlechtert wird.

Die Abhängigkeit der Viskosität  $Q$  vom Datensatz  $U^n$  hat zur Folge, daß das numerische Verfahren, selbst für die lineare Advektionsgleichung, nichtlinear wird. Um die Konservativität des Verfahrens zu erhalten, ist es vorteilhaft, den Viskositätsterm aus (1.45) durch den Ausdruck

$$k [Q(U^n; i)(U_{i+1}^n - U_i^n) - Q(U^n; i-1)(U_i^n - U_{i-1}^n)] \quad (1.46)$$

zu ersetzen. Die künstliche Viskosität  $Q(U^n; i)$  ist jetzt für jedes  $i$  von einer endlichen Menge von Werten aus  $U^n$  abhängig.

In der Praxis ist es sehr schwer, geeignete Viskositäten  $Q$  zu finden, die hinreichende Dissipation in der Umgebung von Unstetigkeiten liefern, ohne die Lösungsstrukturen unnötig stark zu glätten. Da beide Anforderungen gegenläufig sind, kann lediglich ein möglichst guter Kompromiß gefunden werden.

Die in den letzten Jahren entwickelten hochauflösenden Verfahren, wie die TVD- [23], ENO- [25, 62] oder WENO-Methoden [59], basieren auf den verschiedensten Ansätzen, die teilweise zu sehr komplizierten und oft nur implizit definierten Diffusionstermen führen. Die mathematische Untersuchung der resultierenden Verfahren, beispielsweise anhand der modifizierten Gleichungen, ist daher oft sehr aufwendig oder sogar unmöglich.

Im Gegensatz zu diesen hochentwickelten aber rechentechnisch aufwendigen Verfahren betrachten wir in der vorliegenden Arbeit diskrete Filterroutinen aus der Bildverarbeitung, die wir in einfache Differenzenverfahren höherer Ordnung integrieren, um so neue hochauflösende Verfahren für die numerische Lösung hyperbolischer Erhaltungsgleichungen zu konstruieren. Der diskrete Filter läßt sich als nichtlineare Diffusion interpretieren, so daß Hoffnung besteht, auf diese Weise geeignete, sehr einfache und vor allem explizit gegebene künstliche Viskositätsterme zu finden.

Eine Übersicht über verschiedenste existierende hochauflösende Verfahren bieten die Darstellungen von Colella und Woodward [79] und Zalesak [80], grundlegende Beiträge zu der Theorie hochauflösender Methoden sind unter anderem bei Godunov [20], Neumann und Richtmyer [44], Lax [35] und Van Leer [37, 38] zu finden.

## 1.3 Testprobleme

Im Folgenden werden die im Verlauf der vorliegenden Arbeit verwendeten numerischen Testfälle zusammenfassend dargestellt.

### 1.3.1 Eindimensionale Testfälle

#### Lineare Transportgleichung

Im Kontext hyperbolischer Erhaltungsgleichungen steht die Problematik unstetiger Lösungen im Vordergrund. Diese Unstetigkeiten treten nicht nur bei nichtlinearen partiellen Differentialgleichungen auf, sondern sind auch für lineare Erhaltungsgleichungen mit unstetigen Anfangsdaten vorhanden.

Ein klassisches Beispiel für ein lineares Testproblem mit unstetigem Anfangsdatensatz ist das Riemann-Problem für die skalare Advektionsgleichung

$$\begin{aligned} u_t + au_x &= 0, \quad -\infty < x < \infty, \quad t > 0, \\ u_0(x) &= \begin{cases} 1 & : \quad x < 0 \\ 0 & : \quad x > 0 \end{cases}, \end{aligned} \quad (1.47)$$

deren exakte Lösung durch  $u_0(x - at)$  gegeben ist.

#### Burgers' Gleichung

Der bekannteste Testfall für die nichtlineare skalare Erhaltungsgleichung

$$u_t + f(u)_x = 0$$

ist Burgers' Gleichung, für die  $f(u) = \frac{1}{2}u^2$  gilt, so daß sich

$$u_t + uu_x = 0 \quad (1.48)$$

ergibt. Wie bereits in Kapitel 1 diskutiert, erzeugt diese nichtlineare Gleichung selbst für stetige Anfangsbedingungen unstetige Lösungen. In den numerischen Experimenten betrachten wir daher das einfache Cauchy-Problem für Burgers' Gleichung

$$\begin{aligned} u_t + uu_x &= 0, \\ u_0(x) &= \cos(x) + 1.5, \quad x \in [0, 2\pi[, \end{aligned} \quad (1.49)$$

das trotz stetiger Anfangswerte zum Zeitpunkt  $t = 1$  einen Stoß ausbildet.

#### Eindimensionale Eulergleichungen

Die Eulergleichungen beschreiben die Bewegung kompressibler Fluide unter Vernachlässigung der Reibung und resultieren aus der Massen-, Impuls- und Energieerhaltung. Im Eindimensionalen sind die Eulergleichungen, formuliert in konservativen Variablen, gegeben durch

$$\partial_t U + \partial_x f(U) = 0, \quad U := \begin{pmatrix} \rho \\ \rho u \\ \rho E \end{pmatrix}, \quad f(U) := \begin{pmatrix} \rho u \\ \rho u^2 + p \\ \rho u H \end{pmatrix}. \quad (1.50)$$



Dabei bezeichnet  $\rho$  die Dichte des Fluids,  $u$  ist die Flußgeschwindigkeit,  $p$  der Druck,  $E$  die Totalenergie und  $H$  die Enthalpie, definiert durch  $H := E + \frac{p}{\rho}$ . Der Druck ist durch die Zustandsgleichung für ideale Gase,

$$p = (\gamma - 1)\rho(E - \frac{|u|^2}{2}),$$

gegeben, wobei der Isotropenkoeffizient  $\gamma$  unter Annahme von trockener Luft auf 1.4 gesetzt wird.

Existiert die klassische Ableitung  $U_x$ , so läßt sich (1.50) zu

$$U_t + f'(U)U_x = 0 \quad (1.51)$$

umschreiben, und die Jacobi-Matrix  $f'(U)$  ist gegeben durch

$$f'(U) = \begin{pmatrix} 0 & 1 & 0 \\ \frac{\gamma-3}{2}u^2 & (3-\gamma)u & \gamma-1 \\ (\gamma-1)u^3 - \gamma u E & \gamma E - \frac{3}{2}(\gamma-1)u^2 & \gamma u \end{pmatrix}. \quad (1.52)$$

Mit Hilfe der Identität

$$\gamma E = H + \frac{\gamma-1}{2}u^2$$

läßt sich die Jacobi-Matrix mittels

$$f'(U) = \begin{pmatrix} 0 & 1 & 0 \\ \frac{\gamma-3}{2}u^2 & (3-\gamma)u & \gamma-1 \\ \frac{\gamma-1}{2}u^3 - uH & H - (\gamma-1)u^2 & \gamma u \end{pmatrix} \quad (1.53)$$

formulieren.

Das Flußverhalten eines Fluids läßt sich nicht nur in konservativen Variablen, die sich direkt aus den Erhaltungseigenschaften ergeben, beschreiben sondern auch in den primitiven (oder physikalischen) Variablen  $\rho$ ,  $p$  und  $u$ , die als Vektor

$$W = \begin{pmatrix} \rho \\ u \\ p \end{pmatrix}$$

zusammengefaßt werden. Der Vektor der konservativen Variablen ergibt sich dann aus

$$U = \begin{pmatrix} W_1 \\ W_1 W_2 \\ \frac{1}{\gamma-1}W_3 + \frac{1}{2}W_1 W_2^2 \end{pmatrix}.$$

Die Jacobi-Matrix  $Q = \partial U / \partial W$  und ihre Inverse sind gegeben durch

$$Q = \begin{pmatrix} 1 & 0 & 0 \\ u & \rho & 0 \\ \frac{1}{2}u^2 & \rho u & \frac{1}{\gamma-1} \end{pmatrix}, \quad Q^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ -\frac{u}{\rho} & \frac{1}{\rho} & 0 \\ \frac{\gamma-1}{2}u^2 & (1-\gamma)u & \gamma-1 \end{pmatrix}. \quad (1.54)$$

Die Gleichung (1.51) läßt sich gemäß

$$QW_t + f'QW_x = 0$$

umschreiben, und durch Multiplikation mit  $Q^{-1}$  von links ergibt sich die primitive Formulierung der Eulergleichungen zu

$$W_t + \tilde{f}' W_x = 0, \quad \tilde{f}' = Q^{-1} f' Q = \begin{pmatrix} u & \rho & 0 \\ 0 & u & \frac{1}{\rho} \\ 0 & \rho c^2 & u \end{pmatrix}, \quad c^2 = \frac{\gamma p}{\rho}, \quad (1.55)$$

wobei  $c$  die Schallgeschwindigkeit bezeichnet. Die Eigenwerte von  $\tilde{f}'$  (also auch von  $f'$ ) lassen sich damit leicht berechnen und sind durch

$$\lambda_1 = u - c, \quad \lambda_2 = u, \quad \lambda_3 = u + c \quad (1.56)$$

gegeben. Die zugehörigen Eigenvektoren von  $\tilde{f}'$  sind

$$\tilde{R}_1 = \frac{1}{2} \begin{pmatrix} -\rho/c \\ 1 \\ -\rho c \end{pmatrix}, \quad \tilde{R}_2 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \tilde{R}_3 = \frac{1}{2} \begin{pmatrix} \rho/c \\ 1 \\ \rho c \end{pmatrix}.$$

Da die Eigenvektoren linear unabhängig sind, liegt ein hyperbolisches System vor. Zusätzlich sind alle Eigenwerte verschieden, so daß das System strikt hyperbolisch ist.

Für spätere Referenzmöglichkeiten geben wir noch die Eigenvektoren  $R_1$ ,  $R_2$  und  $R_3$  der untransformierten Jacobi-Matrix  $f'(U)$  an. Sie berechnen sich aus  $R_p = Q \tilde{R}_p$ ,  $p = 1, 2, 3$ , und sind gegeben durch

$$R_1 = \frac{1}{2}\rho \begin{pmatrix} -1/c \\ 1 - u/c \\ u - \frac{1}{2}u^2/c - \frac{1}{\gamma-1}c \end{pmatrix}, \quad R_2 = \begin{pmatrix} 1 \\ u \\ \frac{1}{2}u^2 \end{pmatrix}, \quad (1.57)$$

$$R_3 = \frac{1}{2}\rho \begin{pmatrix} 1/c \\ 1 + u/c \\ u + \frac{1}{2}u^2/c + \frac{1}{\gamma-1}c \end{pmatrix}.$$

Weiterführende Theorie der Eulergleichungen und Aspekte ihrer numerischen Lösung finden sich beispielsweise in den Büchern von Hirsch [26, 27] und Wesseling [77].

Anwendung finden die eindimensionalen Eulergleichungen bei den Stoßrohrproblemen der Gasdynamik. Betrachtet wird dabei ein mit Gas gefülltes Rohr, das zunächst durch eine Membran in zwei Bereiche geteilt ist. Das Gas besitzt in einer der Rohrhälften eine höhere Dichte und einen höheren Druck als in der anderen, und die Geschwindigkeit des Gases ist überall Null. Zur Zeit  $t = 0$  wird die Membran vollständig entfernt, und das Gas beginnt, in Richtung des geringeren Drucks zu strömen. Unter der Annahme eines gleichförmigen Flusses durch die Röhre variieren die physikalischen Eigenschaften des Gases lediglich in einer Raumrichtung, und die eindimensionalen Eulergleichungen sind anwendbar.

Als numerische Testfälle werden im Verlauf der Arbeit verschiedene Stoßrohrprobleme aus [69] mit Hilfe eindimensionaler Riemann-Probleme auf dem Intervall  $(0, 1)$  simuliert.

Ein einfaches Stoßrohrproblem ohne sonische Punkte ist durch das Riemann-Problem mit den linken und rechten Zuständen

$$\rho_L = 2.0, \quad u_L = 0, \quad p_L = 2.0, \quad \rho_R = 1.0, \quad u_R = 0, \quad p_R = 1.0 \quad (1.58)$$

gegeben, das an der Stelle  $x_0 = 0.3$  eine Unstetigkeit besitzt. Da die exakte Lösung dieses Stoßrohrproblems berechnet werden kann, ist es ein geeignetes Testproblem für die Bewertung neuer numerischer Verfahren. Nach der Zeit  $t = 0.2$  ergibt sich die in Abbildung 1.1 dargestellte exakte Lösung.

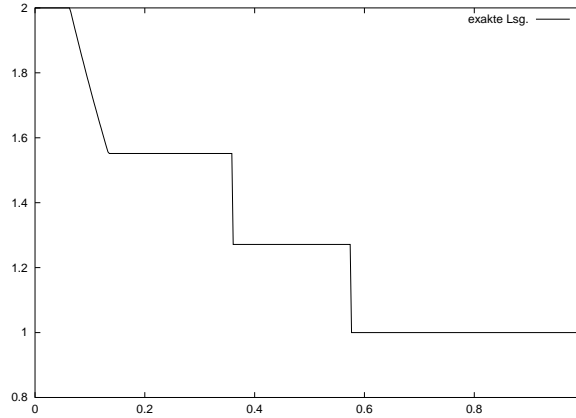


Abbildung 1.1: Lösung des eindimensionalen Riemann-Problems (1.58) zur Zeit  $t = 0.2$

Ein Riemann-Problem mit sonischem Punkt ist durch die Zustände

$$\rho_L = 1.0, u_L = 0.75, p_L = 1.0, \rho_R = 0.125, u_R = 0, p_R = 0.1 \quad (1.59)$$

charakterisiert, wobei die Unstetigkeit wieder bei  $x_0 = 0.3$  liegt. Die exakte Lösung dieses Problems ist für  $t = 0.2$  in Abbildung 1.2 gezeigt.

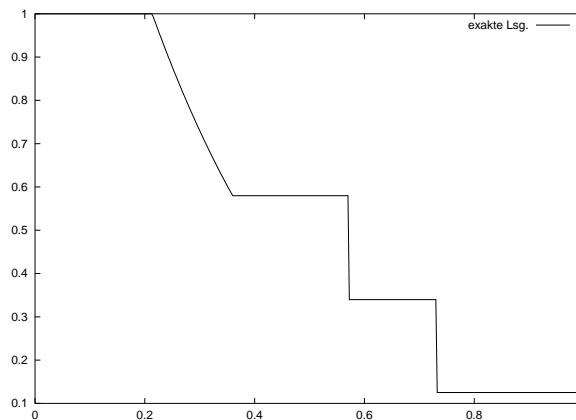


Abbildung 1.2: Lösung des eindimensionalen Riemann-Problems (1.59) zur Zeit  $t = 0.2$

Ein weiteres Riemann-Problem ist der Testfall von Sod, der durch

$$\rho_L = 1.0, u_L = 0, p_L = 1.0, \rho_R = 0.125, u_R = 0, p_R = 0.1 \quad (1.60)$$

beschrieben wird und eine Unstetigkeit bei  $x_0 = 0.5$  aufweist. Die zugehörige exakte Lösung zur Zeit  $t = 0.25$  ist in Abbildung 1.3 dargestellt.

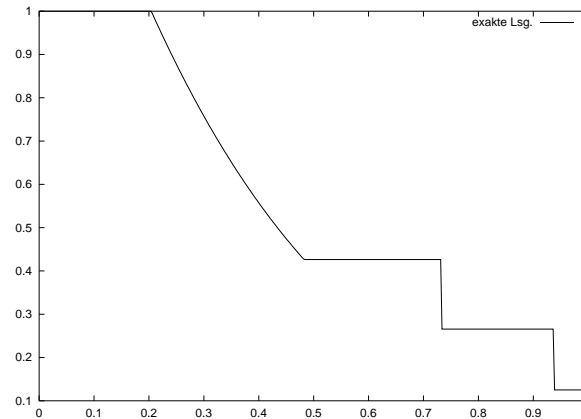


Abbildung 1.3: Lösung des eindimensionalen Riemann-Problems (1.60) zur Zeit  $t = 0.25$

### 1.3.2 Zweidimensionale Testfälle

#### Skalares Testproblem

Als zweidimensionalen, skalaren Testfall betrachten wir das Randwertproblem

$$\partial_t u + \partial_x \frac{u^2}{2} + \partial_y u = 0, \quad (1.61)$$

$$u(x, y, 0) = \begin{cases} 1.5 & , \quad x = 0, \\ -2.5x + 1.5 & , \quad y = 0, \\ -1.0 & , \quad x = 1, \\ 0 & , \quad \text{sonst} \end{cases}.$$

Die Werte am unteren Rand und an den seitlichen Rändern seien fest, und der obere Rand ist ein Ausströmrand.

Die exakte Lösung dieses Anfangswertproblems ist bekannt und besteht aus einer fächerförmigen Welle, die sich zu einer Stoßwelle entwickelt. Abbildung 1.4 zeigt eine schematische Darstellung der stationären Lösung, deren Konstruktion wir kurz erläutern wollen.

Da die exakte Lösung die stationäre Gleichung

$$\partial_y u + \partial_x \frac{u^2}{2} = 0$$

erfüllt, sind die Charakteristiken durch  $dy/ds = 1, dx/ds = u$  gegeben, das heißt,

$$\frac{dy}{dx} = \frac{1}{u}.$$

Bezeichnen wir mit  $u_L$  und  $u_R$  den linken beziehungsweise rechten Zustand bei  $y = 0$  und setzen wir eine lineare Verteilung

$$u(x, 0) = (u_R - u_L)x + u_L$$

der Randwerte bei  $y = 0$  voraus, ist die Gleichung der linken Charakteristik  $g_1$  durch  $y = x/u_L$  gegeben, und die rechte Charakteristik  $g_2$  wird durch  $y = (x - 1)/u_R$  beschrieben. Sie

schneiden sich am Punkt  $P$ , an dem der Stoß  $g_3$  beginnt. Die Koordinaten von  $P$  lassen sich gemäß  $x_P = u_L/(u_L - u_R)$  und  $y_P = 1/(u_L - u_R)$  berechnen. Für den Stoß  $g_3$  ergibt sich aus der Rankine–Hugoniot–Bedingung die Steigung

$$\frac{dy}{dx} = \frac{2}{u_L + u_R}$$

und damit die Gleichung  $y = (2x - 1)/(u_L + u_R)$ . Ist die Lösung an einem Punkt  $Q$ , der innerhalb des fächerförmigen Bereichs liegt, gegeben, schneidet die Charakteristik, die  $P$  und  $Q$  verbindet, die  $x$ -Achse an der Stelle  $x_{PQ} = x_P + y_P(x_Q - x_P)/(y_P - y_Q)$ . Aufgrund der linearen Randverteilung bei  $y = 0$ , die wir vorausgesetzt haben, ergibt sich damit  $u_Q = (u_R - u_L)x_{PQ} + u_L$ , so daß wir die exakte Lösung punktweise berechnen können.

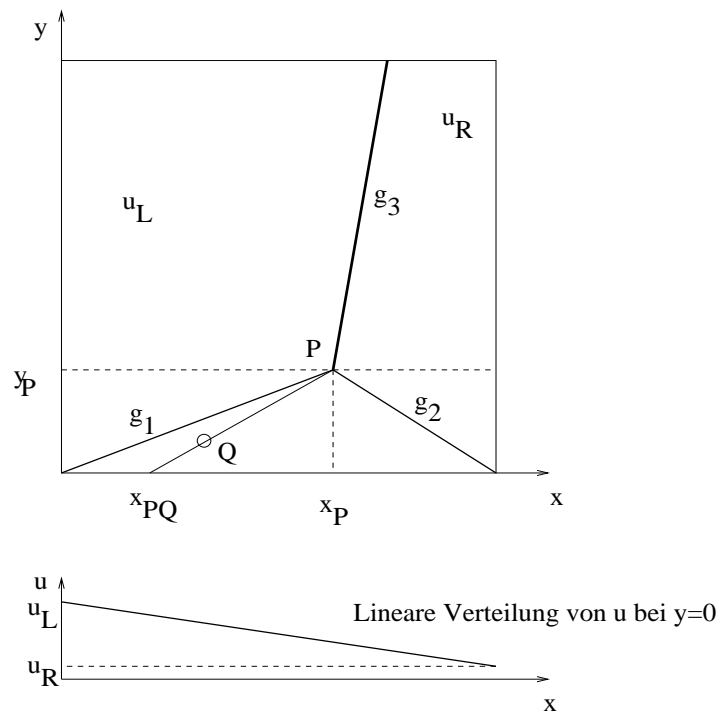


Abbildung 1.4: Exakte Lösung des Testproblems in der  $(x, y)$  Ebene

Da die exakte Lösung dieses Testfalls damit explizit bekannt ist, und sich die Lösung darüber hinaus sowohl aus einem Stoß als auch aus einem stetigen Gebiet zusammensetzt, ist diese Randwertaufgabe ein geeignetes Testproblem für die numerische Lösung zweidimensionaler, skalarer Erhaltungsgleichungen.

## Zweidimensionale Eulergleichungen

Als Beispiel eines zweidimensionalen Systems von Erhaltungsgleichungen betrachten wir die Eulergleichungen der Gasdynamik für zwei Raumdimensionen. Wie bereits in (1.22) darge-

stellt, sind die Eulergleichungen, formuliert in konservativen Variablen, in diesem Fall durch

$$\partial_t U + \partial_x f(U) + \partial_y g(U) = 0 \text{ auf } \mathbb{R}^2 \times \mathbb{R}^+, \quad (1.62)$$

$$U := \begin{pmatrix} \rho \\ \rho u \\ \rho v \\ \rho E \end{pmatrix}, \quad f(U) := \begin{pmatrix} \rho u \\ \rho u^2 + p \\ \rho uv \\ \rho u H \end{pmatrix}, \quad g(U) := \begin{pmatrix} \rho v \\ \rho uv \\ \rho v^2 + p \\ \rho v H \end{pmatrix},$$

gegeben, wobei  $u$  jetzt die Flußgeschwindigkeit in  $x$ -Richtung und  $v$  die Flußgeschwindigkeit in  $y$ -Richtung bezeichnet. Die Enthalpie ist, ebenso wie im Eindimensionalen, durch  $H := E + \frac{p}{\rho}$  definiert. Die Zustandsgleichung für ideale Gase lautet im Zweidimensionalen

$$p = (\gamma - 1) \left[ E - \frac{\rho}{2}(u^2 + v^2) \right],$$

wobei der Isotropenkoeffizient  $\gamma$  unter Annahme von trockener Luft wieder auf 1.4 gesetzt wird.

Für glatte Lösungen kann (1.62) auch in der Form

$$U_t + AU_x + BU_y = 0$$

dargestellt werden, wobei  $A$  und  $B$  die Jacobi-Matrizen der Flußfunktionen  $f$  und  $g$  bezeichnen, die durch

$$A := \nabla_U f(U) = \begin{pmatrix} 0 & 1 & 0 & 0 \\ \frac{\gamma-3}{2}u^2 + \frac{\gamma-1}{2}v^2 & (3-\gamma)u & (1-\gamma)v & \gamma-1 \\ -uv & v & u & 0 \\ -\gamma u E + (\gamma-1)u(u^2 + v^2) & \gamma E - \frac{\gamma-1}{2}(v^2 + 3u^2) & (1-\gamma)uv & \gamma u \end{pmatrix}$$

und

$$B := \nabla_U g(U) = \begin{pmatrix} 0 & 0 & 1 & 0 \\ -uv & v & u & 0 \\ \frac{\gamma-3}{2}v^2 + \frac{\gamma-1}{2}u^2 & (1-\gamma)u & (3-\gamma)v & \gamma-1 \\ -\gamma v E + (\gamma-1)v(u^2 + v^2) & (1-\gamma)uv & \gamma E - \frac{\gamma-1}{2}(u^2 + 3v^2) & \gamma v \end{pmatrix}$$

gegeben sind.

Ebenso wie für die eindimensionalen Eulergleichungen existiert auch hier eine Formulierung in primitiven Variablen, auf die wir an dieser Stelle verzichten wollen. Ausführliche Darstellungen dazu sind in den Büchern von Hirsch [27], Kröner [32] und Wesseling [77] zu finden.

Die Eulergleichungen werden in vielen Gebieten der Aerodynamik eingesetzt, beispielsweise dienen sie zur Modellierung der Luftströmung um ein Flugzeug. Zur Lösung dieses dreidimensionalen Problems sind auch zweidimensionale Fragestellungen, wie die Strömung um ein Tragflächenprofil, das einen Querschnitt eines Flugzeugflügels darstellt, von Interesse. Da schon kleine Änderungen des Tragflächenprofils zu einem ganz anderen Flußverhalten führen können, ist die numerische Simulation solcher Problemstellungen eine wichtige Anwendung der zweidimensionalen Eulergleichungen.

Als Testfall für die vorliegende Arbeit wird die Umströmung des NACA0060 Profils simuliert.

Das Profil wird dabei mit der Mach-Zahl  $Ma = 0.8$  unter dem Winkel  $\alpha = 1.25^\circ$  angeströmt, die Dichte der Anströmung ist  $\rho = 1.0$ . Mit fortschreitender Zeit bildet sich eine stationäre Strömung, die einen starken Stoß auf der Oberseite und einen schwächeren Stoß auf der Unterseite des Profils aufweist. Die als Rechengitter verwendete Triangulierung ist in Abbildung 1.5 dargestellt. Die Diskretisierung beruht auf Finite-Volumen-Methoden für unstrukturierte Gitter, die in [32] und [61] beschrieben werden.

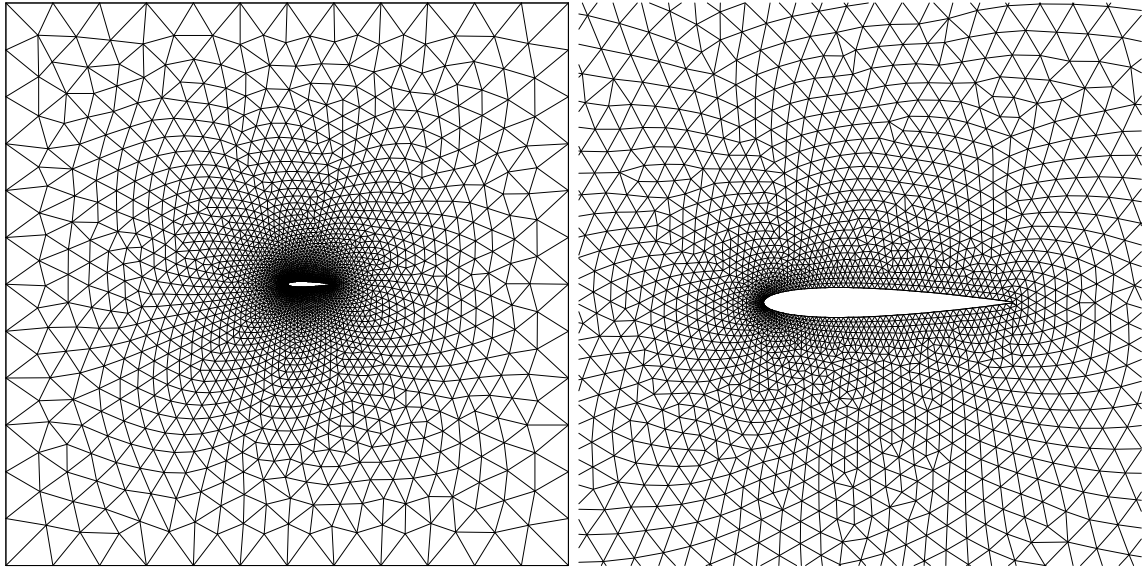


Abbildung 1.5: Triangulierung (rechts ein Ausschnitt) für die Umströmung des NACA0060 Profils





## 2 Grundlegende Herleitungen nichtlinearer Filter

In diesem Kapitel geben wir eine Übersicht über verschiedene existierende Methoden der Bild- beziehungsweise Datenwiederherstellung. Zunächst stellen wir die klassischen Filterkonstruktionen gemäß des Variationsansatzes und verschiedener Modelle partieller Differentialgleichungen vor, die im Kontext der Bildverarbeitung entwickelt worden sind. Bei einer Vielzahl dieser Filtermodelle ist eine Diskrepanz zwischen den Eigenschaften ihrer kontinuierlichen und diskreten Filtergleichungen festzustellen, was die vollständig diskrete Konstruktion von Filteralgorithmen motiviert. Engquist, Lötstedt und Sjögreen folgten dieser Idee und entwickelten 1989 erstmals solche diskreten Filterroutinen, die sie für das Lösen hyperbolischer Erhaltungsgleichungen einsetzten. Wir bestimmen eine zugehörige kontinuierliche Gleichung für einen ihrer diskreten Filter und zeigen damit Verbindungen zu den klassischen Filteralgorithmen der Bildverarbeitung auf.

In den späteren Kapiteln dieser Arbeit werden die drei Varianten der Filterkonstruktion, Variationsansatz, Modelle partieller Differentialgleichungen und vollständig diskrete Herleitungen, immer wieder aufgegriffen, um auf Grundlage dieser Methoden neue nichtlineare Filter zu konstruieren.

### 2.1 Methoden der Bildverarbeitung

Die Entwicklung von Methoden zur Bildwiederherstellung ist eines der ältesten und bedeutendsten Aufgabenfelder der Bildverarbeitung. Auch heute hat diese Fragestellung nicht an Aktualität verloren, da die Bildwiederherstellung bei vielen neuen Anwendungen einen ersten Vorverarbeitungsschritt darstellt.

Bei Übermittlung, Aufzeichnung und Zusammensetzung von Bildern kann die Bildqualität stark beeinträchtigt werden. Einerseits ist dies auf deterministische Phänomene, die durch das Verfahren der Bilderfassung begründet sind, zurückzuführen. Fehler dieser Art entstehen beispielsweise durch Defekte des bilderfassenden Systems oder durch äußere Störungen wie der atmosphärischen Turbulenz. Andererseits wird die Bildqualität bei jeder Signalübermittlung durch eine zufällige Störung, die einem stochastischen Rauschen entspricht, beeinträchtigt. Da die Fehlergrößen nicht als bekannt vorausgesetzt werden können, ist eine Modellvorstellung

für diesen Störvorgang notwendig. Das Zusammenspiel deterministischer und stochastischer Fehlerquellen legt folgendes, oft verwendetes Modell nahe. Sei  $u : \Omega \subset \mathbb{R}^2 \rightarrow \mathbb{R}$  das tatsächliche, unverfälschte und  $u_0$  das beobachtete, gestörte Bild. Dann ist der Zusammenhang von  $u$  und  $u_0$  gegeben durch

$$u_0 = Ru + \eta, \quad (2.1)$$

wobei  $\eta$  das additive weiße Rauschen darstellt und  $R$  ein linearer Operator ist, der die Unschärfe des Bildes verursacht.

Ziel der Bildwiederherstellung ist es, mit Hilfe dieser Modellvorstellung, aus einem beobachteten Bild  $u_0$  eine Approximation an das reale Bild  $u$  zu rekonstruieren.

Im Folgenden geben wir einen Überblick über verschiedene Ansätze zur approximativen Berechnung von  $u$  bei gegebenem  $u_0$ . Dabei orientieren wir uns an dem Buch von Aubert und Kornprobst [5], in dem detailliertere Darstellungen und weitere Methoden der Bildwiederherstellung zu finden sind.

### 2.1.1 Der Variationsansatz

Beim Variationsansatz basiert die Bildwiederherstellung auf der Minimierung eines Funktional, das im Wesentlichen aus zwei Teilen besteht, einem Term, der die Wiedergabetreue zum Ausgangsdatensatz bewertet, sowie einem additiven Regularisierungsterm. Der folgende Abschnitt zeigt die Herleitung dieses Funktional sowie die daraus abgeleiteten Approximationen an das gesuchte Originalbild  $u$ .

Wir setzen voraus, daß zwischen dem Originalbild  $u$  und dem beobachteten Bild  $u_0$  die Beziehung (2.1) gilt. Da  $\eta$  eine Zufallsvariable ist, die Gaußsches weißes Rauschen darstellt, erhalten wir mit der in der Stochastik oft verwendeten Maximum-Likelihood-Methode eine Approximation an  $u$ , indem wir das Kleinste-Quadrate-Problem

$$\inf_u \int_{\Omega} |u_0 - Ru|^2 dx \quad (2.2)$$

lösen, wobei  $\Omega$  der Definitionsbereich des Bildes ist.

Betrachten wir  $u$  und  $u_0$  kurzzeitig als diskrete Variablen des  $\mathbb{R}^m$ ,  $R$  als eine  $m \times m$ -Matrix und  $|\cdot|$  als Euklidische Norm, so folgt aus der Existenz eines Minimums  $u$  von (2.2) die Gültigkeit der Gleichung

$$R^*u_0 - R^*Ru = 0 \quad (2.3)$$

für dieses  $u$ , wobei  $R^*$  der zu  $R$  adjungierte Operator ist. Da  $R^*R$  nicht notwendig eindeutig ist und die zugehörigen Eigenwerte selbst bei Eindeutigkeit sehr klein sein können und zu numerischen Instabilitäten führen, ist das eindeutige Lösen von Gleichung (2.3) nicht möglich. Um dennoch eindeutige Lösungen zu erhalten, betrachten wir im Folgenden regularisierte Versionen des Minimierungsproblems (2.2).

Für die Regularisierung nehmen wir an, daß  $u_0 \in L^\infty(\Omega)$  gilt, wobei  $\Omega \subset \mathbb{R}^2$  beschränkt sei.  $R$  sei ein linearer Operator des  $L^2(\Omega)$ , und  $\eta$  sei eine Störung, die Oszillationen im Bild

hervorrufen.

Wie erstmal 1977 von Tikhonov und Arsenin [68] vorgeschlagen, läßt sich die Regularisierung durch Addition eines Regularisierungsterms vornehmen. Statt des Ausdrucks (2.2) betrachten sie das Minimierungsproblem

$$F(u) = \int_{\Omega} |u_0 - Ru|^2 dx + \lambda \int_{\Omega} |\nabla u|^2 dx, \quad (2.4)$$

wobei der Parameter  $\lambda$  ein positiver Gewichtungsfaktor ist. Der erste Term in  $F(u)$  mißt die Datentreue zum Ausgangsdatensatz, der zweite Ausdruck ist ein Glättungsterm. Die Minimierung von (2.4) bedeutet daher die Suche nach einem  $u$ , das die ursprünglichen Daten bestmöglich approximiert und dabei einen nur kleinen Gradienten aufweist. In dem Funktionenraum

$$W^{1,2}(\Omega) = \{u \in L^2(\Omega); \nabla u \in L^2(\Omega)^2\}$$

sind beide Summanden wohldefiniert und das Problem

$$\inf \{F(u), u \in W^{1,2}(\Omega)\}$$

besitzt eine eindeutige Lösung, die durch die Euler–Lagrange–Gleichung

$$R^*Ru - R^*u_0 - \lambda \Delta u = 0 \quad (2.5)$$

mit Neumannschen Randbedingungen

$$\frac{\partial u}{\partial N} = 0 \text{ auf } \partial\Omega$$

gegeben ist, wobei  $N$  die äußere Normale an  $\partial\Omega$  ist. Allerdings ist es offensichtlich, daß die Lösung  $u$  von (2.5) nicht den gestellten Anforderungen, Oszillationen zu glätten und Kanten zu erhalten, entspricht. Der Laplace–Operator in (2.5) besitzt isotrope Glättungseigenschaften, so daß Kanten überglättet werden. Dieses Verhalten ist bereits an der Energiegleichung (2.4) erkennbar. Die  $L^2$ –Norm des Gradienten entfernt zwar die Oszillationen im beobachteten Bild, reduziert aber gleichzeitig die Gradienten, die den Kanten des Bildes entsprechen, zu stark. Rudin, Osher und Fatemi [52, 54] schlagen daher vor, anstelle der  $L^2$ –Norm die  $L^1$ –Norm des Gradienten von  $u$ , auch Totalvariation von  $u$  genannt, in (2.4) zu verwenden.

Um den Einfluß des Regularisierungsterm genauer zu untersuchen, betrachten wir im Folgenden die Energiegleichung [6, 70]

$$E(u) = \frac{1}{2} \int_{\Omega} |u_0 - Ru|^2 dx + \lambda \int_{\Omega} \phi(|\nabla u|) dx. \quad (2.6)$$

Die Funktion  $\phi$  ist dabei so zu konstruieren, daß die Lösung ein annähernd stückweise konstantes Bild ergibt, das aus homogenen Regionen, getrennt durch scharfe Kanten, besteht. Falls  $E(u)$  ein Minimum  $u$  besitzt, erfüllt es die Euler–Lagrange–Gleichung

$$R^*Ru - \lambda \operatorname{div} \left( \frac{\phi'(|\nabla u|)}{|\nabla u|} \nabla u \right) = R^*u_0. \quad (2.7)$$

Der Divergenzterm läßt sich entsprechend der Bildstrukturen in Richtungen tangential und normal zu den Bildkanten zerlegen. Für jeden Punkt  $x$  mit  $|\nabla u(x)| \neq 0$  definieren wir die

Vektoren  $N(x) = \frac{\nabla u(x)}{|\nabla u(x)|}$  und  $T(x)$ , wobei  $|T(x)| = 1$  gelte und  $T(x)$  orthogonal zu  $N(x)$  sei. Damit ergibt sich für (2.7) die Darstellung

$$R^*Ru - \lambda \operatorname{div} \left( \frac{\phi'(|\nabla u|)}{|\nabla u|} u_{TT} + \phi''(|\nabla u|) u_{NN} \right) = R^*u_0, \quad (2.8)$$

in der  $u_{TT}$  beziehungsweise  $u_{NN}$  die zweiten Ableitungen in Richtung  $T$  beziehungsweise  $N$  bezeichnen.

Mit Hilfe dieser Formulierung läßt sich die gesuchte Funktion  $\phi$  geeignet bestimmen:

- Ein geringer Wert  $|\nabla u|$  weist auf Rauschen hin, das in alle Richtungen gleich geglättet werden soll. Ist  $\phi$  eine reguläre Funktion, so wird isotrope Glättung durch die Forderung

$$\phi'(0) = 0, \quad \lim_{s \rightarrow 0^+} \frac{\phi'(s)}{s} = \lim_{s \rightarrow 0^+} \phi''(s) = \phi''(0) > 0 \quad (2.9)$$

erfüllt, und (2.8) wird zu

$$R^*Ru - \lambda \phi''(0) \Delta u = R^*u_0. \quad (2.10)$$

- Ein großer Wert  $|\nabla u|$  weist auf eine Kante hin, die erhalten werden soll, so daß nur Diffusion entlang der Kante, nicht über sie hinweg, erwünscht ist. Um dies zu gewährleisten, ist der Koeffizient von  $u_{NN}$  in (2.8) für große Gradienten Null zu setzen, wobei der Koeffizient von  $u_{TT}$  nicht verschwinden darf. Die Bedingungen an  $\phi$  lauten daher

$$\lim_{s \rightarrow \infty} \phi''(s) = 0, \quad \lim_{s \rightarrow \infty} \frac{\phi'(s)}{s} = \beta > 0.$$

Da sich diese Forderungen gerade widersprechen, ist ein Kompromiß zwischen beiden Eigenschaften zu finden. Beispielsweise können  $\phi''(s)$  und  $\frac{\phi'(s)}{s}$  jeweils gegen Null konvergieren, allerdings mit unterschiedlicher Konvergenzgeschwindigkeit, so daß

$$\lim_{s \rightarrow \infty} \phi''(s) = \lim_{s \rightarrow \infty} \frac{\phi'(s)}{s} = 0 \quad \text{und} \quad \lim_{s \rightarrow \infty} \frac{\phi''(s)}{\phi'(s)/s} = 0 \quad (2.11)$$

gilt.

Es lassen sich zahlreiche Funktionen  $\phi$  finden, die die Bedingungen (2.9) und (2.11) erfüllen, zum Beispiel

$$\phi(s) = \sqrt{1 + s^2}. \quad (2.12)$$

Allerdings sind diese Bedingungen an  $\phi$  nicht hinreichend für die Existenz einer eindeutigen Lösung für die Minimierung von (2.6).

Für Existenz- und Eindeutigkeitsaussagen bezüglich einer Lösung von

$$\inf \left\{ E(u) = \frac{1}{2} \int_{\Omega} |u_0 - Ru|^2 dx + \lambda \int_{\Omega} \phi(|\nabla u|) dx \right\} \quad (2.13)$$

sind weitere Voraussetzungen, wie beispielsweise die strikte Konvexität von  $\phi$  und die Linearität und Stetigkeit von  $R$ , notwendig. Darüber hinaus muß der betrachtete Funktionenraum

auf  $BV(\Omega)$  erweitert werden. Wir verzichten an dieser Stelle auf die exakte Formulierung der Existenz- und Eindeutigkeitsaussage, die in [2] und [5] zu finden ist, weisen aber auf einen daraus resultierenden Widerspruch zwischen mathematischer Modellvorstellung und numerischer Anwendung hin:

Ohne die Bedingung, daß  $\phi$  eine konvexe Funktion ist, ist weder die Existenz einer Lösung des Minimierungsproblems (2.13) noch die Gültigkeit einer Konvergenzaussage gewährleistet. Dennoch ergeben nichtkonvexe Funktionen, wie

$$\phi(s) = \frac{s^2}{1 + s^2},$$

bei numerischen Anwendungen schärfere Kanten und damit bessere Resultate als konvexe, linear wachsende Funktionen  $\phi$ . Diese widersprüchliche Eigenschaft zwischen diskretem und kontinuierlichem Modell zeigt, daß die Konstruktion von Filtern mit Hilfe kontinuierlicher Modellgleichungen nicht unbedingt zu bestmöglichen numerischen Resultaten führt. Vielmehr stellt sich die Frage nach alternativen Ansätzen zur Filterkonstruktion. In diesem Zusammenhang betrachtet beispielsweise Rosati [51] eine diskrete Version der Energiegleichung  $E(u)$  und untersucht das zugehörige asymptotische Verhalten. Engquist, Lötstedt und Sjögreen gehen noch einen Schritt weiter und lösen sich vollkommen von kontinuierlichen Modellvorstellungen, indem sie allein aus diskreten Überlegungen Filterroutinen zur Wiederherstellung diskreter Datensätze entwickeln. Auf ihre Ansätze wird am Ende dieses Kapitels genauer eingegangen, und ihre Ideen bilden den Hintergrund für viele neue Filtermethoden, die im Verlaufe dieser Arbeit vorgestellt werden.

Zunächst geben wir allerdings eine Übersicht über die Klasse diskreter Filter, die auf partiellen Differentialgleichungen beruhen.

### 2.1.2 Ansätze basierend auf partiellen Differentialgleichungen

Ein weiterer Ansatz für die Entwicklung von Filtermethoden zur Bildrekonstruktion basiert auf der Diskretisierung partieller Differentialgleichungen. In der Literatur finden sich viele verschiedene Modellgleichungen, die sich alle durch die verallgemeinerte Form

$$\begin{aligned} \frac{\partial u}{\partial t}(x, t) + F(x, u(x, t), \nabla u(x, t), \nabla^2 u(x, t)) &= 0 \text{ in } \Omega \times (0, T), \\ \frac{\partial u}{\partial N}(x, t) &= 0 \text{ in } \partial\Omega \times (0, T), \\ u(x, 0) &= u_0(x) \end{aligned} \tag{2.14}$$

darstellen lassen. Dabei bezeichnet  $u(x, t)$  die wiederhergestellte Version des beobachteten, verfälschten Bildes  $u_0(x)$ ,  $\nabla u$  den Gradienten und  $\nabla^2 u$  die Hesse-Matrix von  $u$ , so daß  $F$  ein Differentialoperator zweiter Ordnung ist. Es gelten die Neumannschen Randbedingungen, und  $u(x, 0) = u_0(x)$  sei der gegebene Anfangswert. Im Unterschied zu den vorher betrachteten Gleichungen tritt hier der Parameter  $t$  auf, der in der Bildverarbeitung als Skalierungsvariable interpretiert wird. Ausgehend von  $u_0(x)$  wird bei Anwendung von (2.14) eine Familie von Funktionen und damit Bildern  $\{u(x, t)\}_{t>0}$  so konstruiert, daß sich die Bildstrukturen für wachsendes  $t$  zunehmend vereinfachen, ohne daß neue Strukturen entstehen.

Entscheidend für die Qualität des wiederhergestellten Bildes  $u(x, t)$  ist die Wahl des Operators  $F$ . Dieser Differentialoperator zweiter Ordnung ist so zu wählen, daß  $u(x, t)$  einerseits eine geglättete Version von  $u_0(x)$  ohne Rauschen darstellt, andererseits tatsächliche Bildstrukturen wie Ecken und Kanten von  $u_0(x)$  erhalten bleiben. Um diesen beiden, zunächst widersprüchlich erscheinenden, Forderungen gerecht zu werden, wurden verschiedenste Operatoren  $F$  untersucht.

### Vorwärts parabolische partielle Differentialgleichungen

Die vorwärts parabolischen partiellen Differentialgleichungen bilden bei der Bildwiederherstellung die Klasse der glättenden Filterroutinen. Die älteste und bekannteste Gleichung dieser Kategorie ist die parabolische, lineare Wärmeleitungsgleichung [1, 66]

$$\begin{aligned} \frac{\partial u}{\partial t}(x, t) - \Delta u(x, t) &= 0, \quad x \in \mathbb{R}^2, \quad t > 0, \\ u(x, 0) &= u_0(x). \end{aligned} \quad (2.15)$$

Motiviert wurde dieses Filtermodell durch die Tatsache, daß das Lösen von (2.15) äquivalent zur Anwendung des in der Bildverarbeitung wohlbekannten linearen Gaußschen Filters ist. Die explizite Lösung von (2.15) ist daher gegeben durch

$$u(x, t) = \int_{\mathbb{R}^2} G_{\sqrt{2t}}(x - y) u_0(y) dy = (G_{\sqrt{2t}} * u_0)(x), \quad (2.16)$$

wobei  $G_\sigma(x)$  den zweidimensionalen Gauß-Kern

$$G_\sigma(x) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{|x|^2}{2\sigma^2}\right)$$

bezeichnet. Im Frequenzbereich betrachtet, liefert die Faltung mit einem Gauß-Kern einen Tiefpaßfilter. Das Lösen der Wärmeleitungsgleichung führt damit, unabhängig vom betrachteten Bild, zu einem isotropen Glättungsalgorithmus, der nicht kantenerhaltend ist. Für die Bildwiederherstellung mit Kantenerhaltung ist diese lineare Differentialgleichung daher ungeeignet.

Da es das Ziel ist, abhängig von der Bildstruktur, in verschiedene Bildrichtungen unterschiedlich stark zu glätten, ist es naheliegend, nichtlineare statt lineare Differentialgleichungen zu untersuchen. Perona und Malik [47, 48] betrachten daher als Verallgemeinerung der Wärmeleitungsgleichung das System

$$\begin{aligned} \frac{\partial u}{\partial t} &= \operatorname{div}(c(|\nabla u|^2) \nabla u) \quad \text{in } \Omega \times (0, T), \\ \frac{\partial u}{\partial N} &= 0 \quad \text{auf } \partial\Omega \times (0, T), \\ u(x, 0) &= u_0(x) \quad \text{in } \Omega. \end{aligned} \quad (2.17)$$

Dabei sei  $c(s) : [0, \infty) \rightarrow (0, \infty)$  eine monoton fallende Funktion mit  $c(0) = 1$  und  $\lim_{s \rightarrow \infty} c(s) = 0$ , so daß (2.17) in Regionen mit kleinem Gradienten isotrop glättet, in Regionen mit großen

Gradienten, also in der Umgebung von Kanten, das Glätten hingegen einstellt. Für den Spezialfall  $c \equiv 1$  ergibt sich die Wärmeleitungsgleichung.

Führen wir, wie bereits beim Variationsansatz, Vektoren  $N(x)$  und  $T(x)$  normal beziehungsweise tangential zu den Bildstrukturen ein, läßt sich der Divergenzterm mit Hilfe der Definition

$$b(s) = c(s) + 2c'(s)$$

umformen und für (2.17) ergibt sich die Darstellung

$$\frac{\partial u}{\partial t} = c(|\nabla u|^2)u_{TT} + b(|\nabla u|^2)u_{NN}.$$

Da die Diffusion hauptsächlich entlang einer Kante, nicht über sie hinweg, stattfinden soll, ist  $\lim_{s \rightarrow \infty} \frac{b(s)}{c(s)}$ , also  $\lim_{s \rightarrow \infty} \frac{sc'(s)}{c(s)} = -\frac{1}{2}$ , eine weitere Bedingung an  $c(s) > 0$ , die durch die Eigenschaft

$$\lim_{s \rightarrow \infty} c(s) \approx \frac{1}{\sqrt{s}}$$

gewährleistet wird. Damit (2.17) eine parabolische Differentialgleichung ist, ist darüber hinaus

$$b(s) = c(s) + 2c'(s) > 0$$

zu fordern. Unter diesen Voraussetzungen an  $c(s)$  läßt sich mit Hilfe nichtlinearer Halbgruppentheorie die Existenz einer Lösung von (2.17) beweisen [8, 11].

Da das Modell von Perona und Malik nur den Betrag des Gradienten, nicht seine Orientierung berücksichtigt, untersucht Weickert in [76] eine Erweiterung des Perona–Malik–Modells. Um lokale Variationen der Gradientenorientierung bei der Bildwiederherstellung eingehen zu lassen, betrachtet er die positiv semidefinite Matrix  $\nabla u \nabla u^t$ , die die Eigenwerte  $\lambda_1 = |\nabla u|^2$  und  $\lambda_2 = 0$  besitzt und deren Orthonormalbasis aus den Eigenvektoren  $v_1$ , parallel zu  $\nabla u$ , und  $v_2$ , orthogonal zu  $\nabla u$ , besteht. Um Fehler aufgrund des vorhandenen Rauschens auszuschließen, wird das Bild  $u(x)$  zunächst mit einem Gauß–Kern  $k_\sigma$  gefaltet, so daß sich  $u_\sigma(x) = (k_\sigma * u)(x)$  ergibt. Da auch Informationen aus der Nachbarschaft des Punktes  $x$  berücksichtigt werden sollen, wird auf die Matrix  $\nabla u \nabla u^t$  ebenfalls ein Gauß–Kern  $k_\rho$  angewendet und es ergibt sich die symmetrische, positiv semidefinite Matrix

$$J_\rho(\nabla u_\sigma) = k_\rho * \nabla u_\sigma \nabla u_\sigma^t.$$

Die Eigenvektoren  $v_1, v_2$  von  $J_\rho(\nabla u_\sigma)$  sind orthonormal, wobei  $v_1$  die Richtung maximaler Grauwertschwankungen angibt und  $v_2$  die bevorzugte Glättungsrichtung aufzeigt. Die zugehörigen Eigenwerte  $\mu_1$  und  $\mu_2$  geben Auskunft über die Bildstrukturen. Isotrope Gebiete sind gekennzeichnet durch  $\mu_1 \cong \mu_2$ ,  $\mu_1 \gg \mu_2 \approx 0$  weist auf geradlinige Strukturen und  $\mu_1 \geq \mu_2 \gg 0$  auf Ecken hin.

Damit ergibt sich als sinnvolle Erweiterung von (2.17) die parabolische Gleichung

$$\begin{aligned} \frac{\partial u}{\partial t} &= \operatorname{div}(D(J_\rho(\nabla u_\sigma))\nabla u) \text{ in } \Omega \times (0, T], \\ u(x, 0) &= u_0(x) \text{ auf } \Omega, \\ \langle D(J_\rho(\nabla u_\sigma))\nabla u, N \rangle &= 0 \text{ auf } \partial\Omega \times (0, T], \end{aligned} \tag{2.18}$$

wobei  $N$  die äußere Einheitsnormale an  $\partial\Omega$  bezeichnet und  $D$  ein zu wählender Diffusionstensor ist. Ist  $D$  gleichmäßig positiv definit und gilt  $D \in C^\infty(S^2, S^2)$ , wobei  $S^2$  die Menge der symmetrischen Matrizen bezeichnet, besitzt die parabolische Differentialgleichung (2.18) eine eindeutige Lösung.

Der Diffusionstensor läßt sich abhängig von dem Ziel, das bei der Bildwiederherstellung verfolgt werden soll, geeignet wählen. Da die Eigenvektoren von  $D$  die lokale Bildstruktur wiedergeben sollen, ist es sinnvoll, die Orthonormalbasis von Eigenvektoren von  $D(J_\rho)$  zu übernehmen. Um beim Glätten Kanten zu erhalten, empfiehlt es sich, die Diffusion orthogonal zu Kanten umso stärker zu verringern, je größer  $\mu_1$  ist. Dies kann beispielsweise durch die zugehörigen Eigenwerte

$$\begin{aligned}\lambda_1 &= \begin{cases} 1 & \text{falls } \mu_1 = 0 \\ 1 - \exp\left(\frac{-3.315}{\mu_1^4}\right) & \text{sonst} \end{cases} \\ \lambda_2 &= 1\end{aligned}$$

von  $D(J_\rho)$  gewährleistet werden [75].

Auch wenn natürlich noch weitere vorwärts parabolische Modellgleichungen für die Bildwiederherstellung existieren, sollen die hier vorgestellten Modelle für einen kurzen Einblick genügen, und wir gehen zu entsprechenden rückwärts parabolischen Gleichungen über.

### Rückwärts parabolische partielle Differentialgleichungen

Im vorangehenden Abschnitt wurden glättende Filter betrachtet. Da Bilder oft nicht nur geglättet, sondern Bildstrukturen gleichzeitig verstärkt werden sollen, ist die Idee naheliegend, diese Tiefpaßfilter zusätzlich lokal durch Hochpaßfilter zu ergänzen. Diese Kombination aus glättenden und verstärkenden Filtern läßt sich beispielsweise mit einem geeigneten Gewichtungskoeffizienten  $c(\cdot)$  im Perona–Malik–Modell (2.17) realisieren.

Sei dazu zunächst im Eindimensionalen eine Kante, die durch eine partielle Differentialgleichung verstärkt wird, dadurch gekennzeichnet, daß in einer Umgebung dieser Kante  $\frac{\partial}{\partial t}(u_x(x, t)) \geq 0$  gilt. Entsprechend wird das Verwischen einer Kante durch die Gültigkeit von  $\frac{\partial}{\partial t}(u_x(x, t)) \leq 0$  in einer Umgebung der Kante charakterisiert. Aus dem eindimensionalen Perona–Malik–Modell (2.17) läßt sich

$$\frac{\partial}{\partial t}(u_x) = \left( \frac{\partial u}{\partial t} \right)_x = (c(u_x^2)u_x)_{xx} = u_{xxx}b(u_x^2) + 2u_{xx}^2b'(u_x^2)$$

ableiten, wobei  $b(s) = 2sc'(s) + c(s)$  gilt. Für eine Kante  $x$  zur Zeit  $t$  ist  $u_{xx}(x, t) = 0$  und  $u_{xxx}(x, t) \leq 0$ , so daß

$$\operatorname{sgn} \left( \frac{\partial u}{\partial t}(x, t) \right)_x = \operatorname{sgn} (-b(u_x^2)(x, t))$$

folgt. Es wird deutlich, daß das Verwischen und Verstärken der Kante durch das Vorzeichen von  $b(u_x^2)$  bestimmt wird. Für  $b(u_x^2) > 0$  wird das eindimensionale Perona–Malik–Modell

$$\begin{aligned}\frac{\partial u}{\partial t}(x, t) &= [c(u_x^2(x, t))u_x(x, t)]_x, \\ u(x, 0) &= u_0(x),\end{aligned}\tag{2.19}$$



zu einer vorwärts parabolischen Gleichung und die Kante wird verwischt, für  $b(u_x^2) < 0$  ist (2.19) eine rückwärts parabolische Differentialgleichung und die Kante wird verstärkt.

Diese Veranschaulichung kann auch auf den zweidimensionalen Fall übertragen werden. Für das Verstärken einer Kanten ist zu fordern, daß die Differentialgleichung in Normalenrichtung der Kante rückwärts parabolisch ist, also  $b(s) = 2sc'(s) + c(s) < 0$  für große  $s \geq \lambda$  gilt, wobei  $\lambda$  ein gegebener Schwellenwert ist. Das Glätten homogener Regionen garantiert die Forderung  $c(0) = b(0) = 1$ , denn in diesem Fall ergibt sich für kleine Gradienten die Wärmeleitungsgleichung. Ein klassisches Beispiel für eine mögliche Wahl von  $c(\cdot)$  ist die Funktion

$$c(s) = \frac{1}{1 + \frac{s}{\lambda}}. \quad (2.20)$$

Existenzaussagen über eine Lösung des Perona–Malik–Modells (2.17) mit einem solchen  $c(\cdot)$  lassen sich für den allgemeinen Fall nicht finden. Allerdings konnten Catté und Mitarbeiter [10] für bestimmte Fälle die Existenz einer Lösung im maßtheoretischen Sinn für eine regularisierte Version des Modells zeigen.

Obwohl mathematische Aussagen über die Existenz einer Lösung des Perona–Malik–Modells (2.17) fehlen, wird es bei der Bildwiederherstellung oft verwendet und liefert sehr gute numerische Ergebnisse.

## Hyperbolische partielle Differentialgleichungen

Die letzte Klasse partieller Differentialgleichungen, die wir im Folgenden betrachten, ist die Klasse der hyperbolischen Differentialgleichungen. Sie werden bei der Bildwiederherstellung eingesetzt, um überglättete Kanten wieder aufzuteilen. Diese Kantenverschärfung läßt sich als gegenteiliger Prozeß zur Kantenglättung auffassen. Gegeben ist eine geglättete Kante  $u_0(x)$ , aus der mit Hilfe einer hyperbolischen partiellen Differentialgleichung die tatsächliche Kante  $u(x)$  rekonstruiert werden soll.  $u_0(x)$  sei dabei durch die Anfangsbedingung  $u_0(x) = \cos(x)$  gegeben und die optimale Kante  $u(x)$  läßt sich mit Hilfe der Treppenfunktion

$$u(x) = \begin{cases} 1 & \text{falls } x > 0 \\ -1 & \text{falls } x < 0 \end{cases}$$

beschreiben. Anschauliche Überlegungen zeigen, daß bei dem gewünschten Übergang von  $u_0(x)$  zu  $u(x)$ , der durch eine Familie von Funktionen  $\{u(x, t)\}_{t>0}$  beschrieben werden kann, die Punkte  $x$  mit  $u_{xx}(x, t) = 0$  oder  $u_x(x, t) = 0$  konstant bleiben sollen, während für alle anderen Punkte  $x$  eine Bewegung abhängig vom Vorzeichen von  $u_x(x, t)u_{xx}(x, t)$  wünschenswert wäre.

Aufgrund dieser Überlegungen entwickeln Rudin und Osher [46] die hyperbolische partielle Differentialgleichung

$$\begin{aligned} u_t(x, t) &= -|u_x(x, t)| \operatorname{sgn}(u_{xx}(x, t)), \\ u(x, 0) &= u_0(x), \end{aligned} \quad (2.21)$$

wobei  $\operatorname{sgn}(0) = 0$  gelte. Diese nichtlineare Gleichung liefert das geforderte Verhalten, beispielsweise ergibt sie für Punkte  $x$  mit  $u_x(x, t) > 0$  und  $u_{xx}(x, t) > 0$  eine Transportgleichung mit der Geschwindigkeit eins.

Rudin und Osher verallgemeinern die Differentialgleichung (2.21) weiter und betrachten das auch als Stoßfilter bezeichnete Modell

$$\begin{aligned} u_t(x, t) &= -|u_x|F(u_{xx}), \quad x \in \mathbb{R}, t > 0, \\ u(x, 0) &= u_0(x). \end{aligned} \quad (2.22)$$

Die Funktion  $F$  sei dabei Lipschitz-stetig und erfülle die Eigenschaften  $F(0) = 0$  und  $\operatorname{sgn}(s)F(s) > 0$  für  $s \neq 0$ . Ein klassisches Beispiel für  $F$  ist  $F(u_{xx}) = u_{xx}$ , so daß (2.22)

$$\begin{aligned} u_t + (u_{xx}\operatorname{sgn}(u_x))u_x &= 0, \quad x \in \mathbb{R}, t > 0, \\ u(x, 0) &= u_0(x) \end{aligned} \quad (2.23)$$

ergibt. Diese Gleichung kann als Transportgleichung aufgefaßt werden, deren Ausbreitungsgeschwindigkeit lokal durch  $c(x) = \operatorname{sgn}(u_x)u_{xx}$  gegeben ist und als Kantendetektor interpretiert werden kann, da  $c(x)$  beim Auftreten einer Kante Null wird und  $u_{xx}$  in einer Umgebung der Kante das Vorzeichen wechselt.

Die Übertragung der Modellgleichungen auf den zweidimensionalen Fall verläuft ohne Schwierigkeiten. Statt (2.22) betrachten Rudin und Osher im Zweidimensionalen den Stoßfilter

$$\frac{\partial u}{\partial t} = -|\nabla u|F(L(u)), \quad (2.24)$$

wobei  $F$  wieder  $F(0) = 0$  und  $\operatorname{sgn}(s)F(s) > 0$  für  $s \neq 0$  erfülle und  $L$  ein nichtlinearer elliptischer Operator zweiter Ordnung sei, dessen Nullstellen gerade die Kanten des Bildes beschreiben. Ein klassisches Beispiel für  $L$  ist der Laplace-Operator

$$L(u) = \Delta u = u_{xx} + u_{yy},$$

bessere Ergebnisse lassen sich allerdings mit

$$L(u) = \frac{1}{|\nabla u|^2}(u_x^2 u_{xx} + 2u_x u_y u_{xy} + u_y^2 u_{yy}),$$

der zweiten Ableitung von  $u$  in Richtung  $\frac{\nabla u}{|\nabla u|}$ , erzielen.

Wie bei vielen anderen vorangegangenen kontinuierlichen Differentialgleichungen existiert auch hier, schon im Eindimensionalen, keine eindeutige Lösung von (2.23). Dennoch konnten Rudin und Osher mit diesen Stoßfiltern sehr gute numerische Ergebnisse erzielen, was auf die Diskrepanz zwischen Eigenschaften kontinuierlicher und diskreter Filtergleichungen hinweist.

## 2.2 Vollständig diskrete Filterroutinen

Aufgrund dieser Unterschiede zwischen kontinuierlichem und diskretem Modell entwickelte sich um 1989 ein weiterer Ansatz zur Konstruktion diskreter Filter. Entgegen den bisher beschriebenen Filtern, die auf Modellgleichungen partieller Differentialgleichungen oder der Minimierung einer Energiegleichung beruhen, basiert dieser neue Ansatz darauf, jegliche Filtermotivierung durch kontinuierliche Modellgleichungen außer Acht zu lassen und Filteralgorithmen lediglich aufgrund des gegebenen diskreten Datensatzes zu entwickeln. Die Ersten, die mit Hilfe dieser vollständig diskreten Methode Filterroutinen, nicht im Kontext

der Bildverarbeitung sondern für das numerische Lösen hyperbolischer Erhaltungsgleichungen, konstruiert haben, sind Engquist, Lötstedt und Sjögreen [16]. Ihre Filter werden wir im Folgenden vorstellen und erklären. Darüber hinaus entwickeln wir eine zugehörige kontinuierliche Filtergleichung, um Verwandtschaften mit den bisher beschriebenen Filtermethoden der Bildverarbeitung aufzuzeigen.

### 2.2.1 Erste diskrete Filteralgorithmen von Engquist, Lötstedt und Sjögreen

In ihrem Artikel *Nonlinear Filters for Efficient Shock Computation* [16] veröffentlichten Engquist, Lötstedt und Sjögreen 1989 erstmals vollständig diskrete Filteroperatoren für die numerische Approximation hyperbolischer Erhaltungsgleichungen mit unstetigen Lösungen.

Als Ausgangssituation betrachten sie vorerst im Eindimensionalen ein System nichtlinearer Erhaltungsgleichungen

$$\begin{aligned}\partial_t u + \partial_x f(u) &= 0, \\ u(x, 0) &= u_0(x)\end{aligned}\tag{2.25}$$

mit  $u : \mathbb{R} \times \mathbb{R}_+ \ni x, t \mapsto u(x, t) \in D \subset \mathbb{R}$  und Flußfunktion  $f$ . Wie bereits verdeutlicht, sind Gleichungen dieser Art dafür bekannt, daß sie Unstetigkeiten hervorbringen, selbst wenn die Anfangswertbedingung  $u_0(x)$  beliebig glatt ist. Für die numerische Approximation einer schwachen Lösung von (2.25) wird zunächst ein bereits bekanntes, konsistentes Finite-Differenzen-Schema  $G$  angewendet, so daß sich

$$\begin{aligned}U_i^{n+1} &= G(U_{i-r}^n, U_{i-r+1}^n, \dots, U_{i+r}^n), \\ U_i^0 &= U_0(x),\end{aligned}$$

kurz  $U^{n+1} = G(U^n)$ , mit  $n = 0, 1, 2, \dots$ ,  $i = \dots, -1, 0, 1, \dots$ ,  $x_i = i\Delta x$ ,  $t_n = n\Delta t$  und  $\Delta t = \lambda\Delta x$  ergibt. Das Differenzenverfahren  $G$  sei dabei ein Verfahren höherer Ordnung, so daß bei seiner Anwendung Oszillationen in der Lösung  $U^{n+1}$  entstehen. Um diese Oszillationen zu beseitigen, wird nach jedem Zeitschritt des Differenzenverfahrens eine nichtlineare Filterroutine  $P$  eingefügt und es ergibt sich eine gefilterte Lösung  $\tilde{U}^{n+1}$  gemäß

$$\tilde{U}^{n+1} = P(U^{n+1}, U^n).$$

Für die sinnvolle Filterkonstruktion sind folgende natürliche Bedingungen von  $P$  zu erfüllen:

- *Konsistenz*: In glatten Regionen soll der Filter die Approximation  $U^{n+1}$  möglichst wenig verändern, so daß

$$\|\tilde{U}^{n+1} - U^{n+1}\| = O(\Delta t^2) \text{ für glattes } U^{n+1}\tag{2.26}$$

für ein Verfahren zweiter Ordnung zu fordern ist.

- *Erhaltungsform*: Um die Konvergenz gegen die richtige schwache Lösung zu gewährleisten, muß für das zusammengesetzte Verfahren

$$\sum_i \varphi(x_i)(\tilde{U}_i^{n+1} - U_i^{n+1}) \leq C\tag{2.27}$$

gelten, wobei  $\varphi$  eine geeignete Testfunktion sei.

- *TVD*: Damit die Oszillationen in der Nähe der Unstetigkeiten verringert werden, ist die Ungleichung

$$\begin{aligned} TV(\tilde{U}^{n+1}) &\leq TV(\tilde{U}^n), \\ TV(\tilde{U}^n) &= \sum_i |\tilde{U}_{i+1}^n - \tilde{U}_i^n| \end{aligned} \quad (2.28)$$

zu erfüllen.

- *Minimale Änderung*: Der Filterschritt soll Eigenschaft (2.28) mit möglichst wenigen Operationen erfüllen, ohne dabei (2.26) oder (2.27) zu verletzen. Das bedeutet, daß

$$|\tilde{U}_i^{n+1} - U_i^{n+1}| \neq 0 \quad (2.29)$$

für so wenige Werte von  $i$ , wie möglich, gelte.

Da die Filterroutinen von Engquist, Lötstedt und Sjögreen gleichzeitig möglichst einfach sein sollen, erfüllen nicht alle ihre Filter die Bedingungen (2.28) und (2.29) vollständig.

Als einfachsten Filter für die Situation, daß der zu filternde Datensatz  $U^{n+1}$  nach einem Schritt des Differenzen-Verfahrens  $G$  durch  $U^{n+1} = (U_1^{n+1}, U_2^{n+1}, \dots, U_N^{n+1})^T$  gegeben ist, schlagen Engquist, Lötstedt und Sjögreen folgenden Algorithmus vor:

```

for  $i = 2$  to  $N - 1$  do
    if  $(\Delta_+ U_i)(\Delta_- U_i) < 0$  then
        if  $|\Delta_+ U_i| > |\Delta_- U_i|$  then
             $\delta_+ := |\Delta_+ U_i|$ 
             $\delta_- := |\Delta_- U_i|$ 
             $i_{\text{Corr}} := i + 1$ 
        else
             $\delta_+ := |\Delta_- U_i|$ 
             $\delta_- := |\Delta_+ U_i|$ 
             $i_{\text{Corr}} := i - 1$ 
        endif
         $\delta := \min\{\delta_-, \frac{\delta_+}{2}\}$ 
         $s := \text{sgn}(\Delta_+ U_i)$ 
         $U_i := U_i + s\delta$ 
         $U_{i_{\text{Corr}}} := U_{i_{\text{Corr}}} - s\delta$ 
    endif
endfor

```

wobei  $\Delta_+ U_i = U_{i+1} - U_i$  und  $\Delta_- U_i = -(U_{i-1} - U_i)$  gelte und zur Vereinfachung der Schreibweise der Index  $n + 1$  weggelassen wurde.

Befindet sich beispielsweise ein Extremwert im Punkt  $i$  und gilt  $|\Delta_- U_i| > |\Delta_+ U_i|$ , so ist der Filter gegeben durch

$$\tilde{U}_i = U_i + \operatorname{sgn}(\Delta_+ U_i) \cdot \min \left( |\Delta_+ U_i|, \frac{|\Delta_- U_i|}{2} \right). \quad (2.31)$$

Die wesentlichen Schritte des Filters bestehen aus dem Auffinden der Extremwerte von  $U^{n+1}$  und ihrer Verminderung beziehungsweise Erhöhung, je nachdem, ob ein Maximum oder Minimum vorliegt. Der Filter wird nur an wenigen Gitterpunkten aktiv, und alle Korrekturen werden lokal und in konservativer Weise durchgeführt. Genauer betrachtet, arbeitet die Filterroutine wie folgt:

- Der Filter läuft über die Funktionswerte  $U_i$ ,  $i = 1, \dots, N$  und korrigiert die Werte  $U_i$ , die lokale Minima oder Maxima sind.
- Die Korrektur besteht aus einer Vergrößerung der Minima und einer Verkleinerung der Maxima.
- Bei der Addition einer Korrektur zu einem Punkt  $U_i$ , wird die gleiche Korrektur von einem der direkten Nachbarpunkte abgezogen, wobei der Nachbar ausgewählt wird, dessen Funktionswert die größte Differenz zum Funktionswert des Punktes  $U_i$  aufweist.
- Kein Punkt wird so stark korrigiert, daß sein Funktionswert einen Funktionswert seiner Nachbarpunkte passiert, was gewährleistet, daß keine neuen Extremwerte entstehen.

Obwohl dieser einfache Filter für numerische Anwendungen bereits gute Ergebnisse liefert, besitzt er einige Defizite. Er erkennt keine Extremwerte, die aus mehr als einem Punkt bestehen, die Totalvariation kann ansteigen und er glättet nicht nur durch Oszillationen hervorgerufene Extremwerte. Um diese unerwünschten Eigenschaften zu beheben, stellen Engquist, Lötstedt und Sjögreen in [16] noch eine Reihe weiterer, technisch ausgefeilterer Algorithmen vor.

Da es uns die Einfachheit ihres ersten Filters ermöglicht, eine zugehörige kontinuierliche Form für die diskrete Filterroutine zu entwickeln, bleiben wir bei dem oben beschriebenen Algorithmus (2.30) und leiten im Folgenden eine kontinuierliche Modellgleichung her.

Wir betrachten die Situation, daß ein Extremwert im Punkt  $i$  detektiert wurde und  $|\Delta_- U_i| > |\Delta_+ U_i|$  gelte. Für diesen Fall ist der Filter von Engquist, Lötstedt und Sjögreen durch (2.31) gegeben. Um diesen diskreten Filter auf eine kontinuierliche Form zu bringen, führen wir zunächst eine künstliche Zeitskala  $\Delta\tau$  ein und erhalten

$$\frac{\tilde{U}_i - U_i}{\Delta\tau} = \frac{1}{\Delta\tau} \operatorname{sgn}(\Delta_+ U_i) \cdot \min \left( |\Delta_+ U_i|, \frac{|\Delta_- U_i|}{2} \right), \quad (2.32)$$

so daß die linke Seite der Gleichung bereits den Ausdruck  $\partial_\tau u|_i$  ergibt.

Die Terme  $|\Delta_+ U_i|$  und  $|\Delta_- U_i|$  werden mit Hilfe der Taylorentwicklung umgeformt zu

$$\begin{aligned} |\Delta_+ U_i| &= |U_{i+1} - U_i| \approx |u(x_{i+1}) - u(x_i)| = |u(x+h) - u(x)| \\ &= \left| hu'(x) + \frac{h^2}{2}u''(x) + O(h^3) \right| \leq |u'(x)|h + \left| \frac{h^2}{2}u''(x) + O(h^3) \right| \end{aligned}$$

und analog

$$|\Delta_- U_i| \leq |u'(x)|h + \left| -\frac{h^2}{2}u''(x) + O(h^3) \right|.$$

Damit läßt sich das Minimum in (2.32) abschätzen zu:

$$\begin{aligned} & \min \left( |\Delta_+ U_i|, \frac{|\Delta_- U_i|}{2} \right) \\ & \leq \min \left( |u'(x)|h + \left| \frac{h^2}{2} u''(x) + O(h^3) \right|, \left| u'(x) \frac{h}{2} + \frac{1}{2} \left| -\frac{h^2}{2} u''(x) + O(h^3) \right| \right| \right). \end{aligned}$$

Vernachlässigen wir Terme der Größenordnung  $h^2$  und kleiner, so ergibt sich

$$\begin{aligned} \min \left( |\Delta_+ U_i|, \frac{|\Delta_- U_i|}{2} \right) &= |u'(x)| \cdot \min \left( h, \frac{h}{2} \right) + O(h^2) \\ &= |\nabla_x u| \cdot \frac{h}{2} + O(h^2) \quad (h > 0) \end{aligned}$$

und

$$\text{sgn}(\Delta_+ U_i) = \text{sgn}(\nabla_x u \cdot h + O(h^2)).$$

Mit der Bezeichnung  $F(\cdot) = \frac{h}{2\Delta\tau} \text{sgn}(\cdot)$  und der Annahme  $\text{sgn}(\Delta_+ U_i) = \text{sgn}(\nabla_x u \cdot h)$  ergibt sich

**Lemma 2.1**

*Der in (2.31) gegebene diskrete Filter besitzt mit obigen Bezeichnungen in erster Näherung die kontinuierliche Form*

$$\begin{aligned} \partial_\tau u &= |\nabla_x u| \frac{h}{2\Delta\tau} \text{sgn}(\nabla_x u \cdot h) \\ &= |\nabla_x u| F(\nabla_x u \cdot h). \end{aligned} \tag{2.33}$$

Bei genauerer Betrachtung von (2.33) ergibt sich sofort

**Proposition 1**

*An der kontinuierlichen Form (2.33) des Filters läßt sich direkt das tatsächliche Filterverhalten für die verschiedenen diskreten Situationen der Filteraktivierung ablesen.*

**Beweis**  $|\nabla_x u| \cdot \frac{h}{2}$  gibt den Betrag an, um den der Filter in der Zeit  $\Delta\tau$  korrigiert. Für positives  $\nabla_x u$  (dies ist der Fall, wenn der Wert von  $u$  am folgenden Gitterpunkt größer ist als am gerade betrachteten Punkt) liefert die Signumfunktion ein positives Vorzeichen und der Korrekturbetrag wird addiert, wie man es für einen zu niedrigen Wert von  $u$  erwartet. Die Signumfunktion liefert dagegen ein negatives Vorzeichen, wenn  $u$  an der betrachteten Stelle verglichen mit dem Wert am folgenden Gitterpunkt zu groß ist, so daß die Korrektur in diesem Falle, wie erwartet, subtrahiert wird, jeweils vorausgesetzt, daß an der betrachteten Stelle vorab ein Extremum festgestellt wurde. ■

Lemma 2.2 verdeutlicht, daß die kontinuierliche Version (2.33) der diskreten Filterroutine von Engquist, Lötstedt und Sjögreen für  $u_x > 0$  gerade eine Transportgleichung mit der Geschwindigkeit  $-\frac{h}{2\Delta\tau}$  ist, für  $u_x < 0$  besitzt sie die Geschwindigkeit  $\frac{h}{2\Delta\tau}$ .

Die kontinuierliche Filtergleichung (2.33) ermöglicht uns darüber hinaus einen Vergleich des Filters mit den Filtermethoden der Bildverarbeitung, wie der folgende Abschnitt zeigt.

### 2.2.2 Zusammenhang zu Methoden der Bildverarbeitung

Obwohl sich die Konstruktion diskreter Filter für numerische Lösungen hyperbolischer Erhaltungsgleichungen auf dem ersten Blick von dem Ziel der Bildrekonstruktion, das in der Bildverarbeitung angestrebt wird, unterscheidet, zeigt eine genauere Betrachtung, daß bei beiden Anwendungen sehr ähnliche Absichten verfolgt werden.

Beim numerischen Lösen nichtlinearer, hyperbolischer Erhaltungsgleichungen mit Hilfe Finiten-Differenzen-Verfahren höherer Ordnung entstehen, bedingt durch das Differenzenverfahren, unphysikalische Oszillationen in der Umgebung von Unstetigkeiten. Diese Oszillationen, die die Lösung verfälschen, sollen nachträglich durch nichtlineare Filterroutinen entfernt werden, ohne dabei die Unstetigkeiten selbst zu zerstören. Interpretieren wir die Unstetigkeitsstellen einer numerischen Lösung als Kanten oder Singularitäten eines Bildes und die Oszillationen in der Lösung als Kombination aus stochastischen und deterministischen Störungen dieses Bildes, entspricht das Glätten der Oszillationen bei Erhaltung der Unstetigkeiten gerade der Bildrekonstruktion, bei der durch Übertragungsfehler verfälschte Bilder wiederhergestellt werden sollen, ohne dabei die eigentlichen Bildstrukturen zu beeinträchtigen.

Um die Verwandtschaft der, für die jeweiligen Anwendungen, entwickelten Filterroutinen zu dokumentieren, vergleichen wir die kontinuierliche Gleichung, die wir für den diskreten Algorithmus von Engquist, Lötstedt und Sjögreen hergeleitet haben, mit den Modellen der Bildverarbeitung.

Zunächst läßt sich die kontinuierliche Form (2.33) des diskreten Filters umformen zu

$$\begin{aligned}\partial_\tau u &= |u_x| \frac{h}{2\Delta\tau} \operatorname{sgn}(u_x \cdot h) \\ &= |u_x| \frac{h}{2\Delta\tau} \operatorname{sgn}(u_x) \\ &= |u_x| F(u_x),\end{aligned}\tag{2.34}$$

wobei  $F(u_x) = \frac{h}{2\Delta\tau} \operatorname{sgn}(u_x)$  gilt, und eine Übertragung ins Zweidimensionale liefert

$$\begin{aligned}\partial_\tau u &= |\nabla u| \frac{h}{2\Delta\tau} \operatorname{sgn}(\nabla u) \\ &= |\nabla u| F(\nabla u)\end{aligned}\tag{2.35}$$

mit  $F(\nabla u) = \frac{h}{2\Delta\tau} \operatorname{sgn}(\nabla u)$ .

Betrachten wir die in Kapitel 2.1.2 vorgestellten Algorithmen der Bildverarbeitung, fällt die Ähnlichkeit von Gleichung (2.34) beziehungsweise (2.35) zu den Stoßfiltern von Rudin und Osher auf, die im Eindimensionalen durch

$$\begin{aligned}u_t(x, t) &= -|u_x| F(u_{xx}), \quad x \in \mathbb{R}, \quad t > 0, \\ u(x, 0) &= u_0(x)\end{aligned}\tag{2.36}$$

und im Zweidimensionalen durch

$$\frac{\partial u}{\partial t} = -|\nabla u| F(L(u))\tag{2.37}$$

gegeben sind, wobei die Funktion  $F$  die Eigenschaften  $F(0) = 0$  und  $\operatorname{sgn}(s)F(s) > 0$  für  $s \neq 0$  erfüllt und  $L$  ein nichtlinearer elliptischer Operator zweiter Ordnung ist. Seine Nullstellen entsprechen gerade den Kanten des Bildes, da  $L(0) = 0$  bedeutet, daß die zweiten Ableitungen verschwinden und damit die Krümmung Null ist, was ein Merkmal für eine Kante ist. Als Anfangswert wird das verrauschte Signal  $u_0(x)$  verwendet.

Bei einem Vergleich der Modellgleichungen (2.35) und (2.37) wird deutlich, daß sich die beiden Filteranweisungen, abgesehen vom unterschiedlichen Vorzeichen, in dem Operator  $L(u)$  unterscheiden. Für die Klasse der Stoßfilter stellt  $L(u)$  einen elliptischen Operator zweiter Ordnung dar, in (2.35) entspricht  $L(u)$  dagegen dem Ausdruck  $\nabla u$ , also einem Operator erster Ordnung. Entsprechendes gilt für den eindimensionalen Fall, im Stoßfilter (2.36) von Rudin und Osher ist  $F$  eine Funktion von  $u_{xx}$ , in (2.34) ist  $F$  dagegen abhängig von  $u_x$ . Dieser Unterschied zwischen den Modellgleichungen findet sich in den Eigenschaften der diskreten Filter wieder. Während die diskrete Version des Stoßfilters (2.36) Totalvariationsvermindernd ist, kann die Anwendung der diskreten Filterroutine von Engquist, Lötstedt und Sjögreen die Totalvariation erhöhen. Ein Unterschied zwischen ihren kontinuierlichen Modellen ist daher nur natürlich.

Es stellt sich die Frage, ob sich die Information aus der Bildverarbeitung, daß diskrete Versionen von (2.36) Totalvariationsvermindernd sind, dazu verwenden läßt, die durch (2.30) gegebene Filterroutine in ein TVD-Verfahren umzuformen.

Im vorliegenden Fall ist es dazu erforderlich, den Operator  $L(u)$ , der in der kontinuierlichen Form (2.33) durch  $\nabla_x u \cdot h$  gegeben ist, durch einen Operator zweiter Ordnung zu ersetzen. Wir wählen zu diesem Zweck den Laplace-Operator  $\Delta u$  multipliziert mit  $h$ . Gesucht ist damit ein diskreter Ausdruck, dessen kontinuierliche Form näherungsweise die Darstellung  $\Delta u \cdot h$  besitzt. Mit Hilfe der Taylorentwicklung läßt sich zeigen, daß beispielsweise  $\Delta_+ U'_i$  diese Bedingung erfüllt, denn:

$$\begin{aligned} \Delta_+ U'_i &= U'_{i+1} - U'_i \approx u'(x_{i+1}) - u'(x_i) = u'(x+h) - u'(x) \\ &= u'(x) + hu''(x) + \frac{h^2}{2}u'''(x) + O(h^3) - u'(x) \\ &= hu''(x) + O(h^2) = \Delta u \cdot h + O(h^2) . \end{aligned}$$

Unter Berücksichtigung des negativen Vorzeichens aus (2.36) erhalten wir damit den neuen diskreten Filteralgorithmus

$$\tilde{U}_i = U_i - \operatorname{sgn}(\Delta_+ U'_i) \cdot \min \left( |\Delta_+ U_i|, \frac{|\Delta_- U_i|}{2} \right), \quad (2.38)$$

wobei noch eine geeignete Diskretisierung von  $\Delta_+ U'_i$  zu wählen ist.

Im Gegensatz zum ursprünglichen Algorithmus (2.31) ist die Signumfunktion nun nicht mehr von  $U_i$  und  $U_{i+1}$  abhängig, sondern von deren Ableitungen und damit von den Werten  $U_{i-1}$ ,  $U_i$ ,  $U_{i+1}$  und  $U_{i+2}$ .

Um aus (2.38) einen sinnvollen Filteralgorithmus zu erhalten, muß die Einschränkung gemacht werden, daß der zu filternde Datensatz keine aufeinander folgenden Extremwerte beinhalten darf, auch nicht in Form von Extremwertplateaus, denn dann wäre  $\Delta_+ U'_i = 0$  und der Filter



würde wegen  $\text{sgn}(0) = 0$  nichts bewirken.

Auffällig ist, daß Engquist, Lötstedt und Sjögreen für ihre technisch ausgereifteren Filteralgorithmen [16] genau diese Einschränkung, das Verbot zweier aufeinanderfolgender Extremwerte, machen, beziehungsweise diese Situation getrennt behandeln.

Dies legt die Vermutung nahe, daß der oben geschilderte Versuch, ausgehend von der kontinuierlichen Form TV-vermindernder Filter aus der Bildverarbeitung, diskrete Filter mit gleicher Eigenschaft für die Numerik hyperbolischer Erhaltungsgleichungen zu entwickeln, sinnvoll ist, wenngleich der auf diese Weise hergeleitete Filteralgorithmus (2.38) aufgrund der verwendeten Näherungen nicht notwendig TV-vermindernd ist.

Insgesamt wird deutlich, daß ein Zusammenhang zwischen den diskreten Filtern, die für das Lösen hyperbolischer Erhaltungsgleichungen entwickelt wurden, und denen, die für die Bildrekonstruktion eingesetzt werden, besteht. Diese Tatsache kann nicht nur seitens der Bildverarbeitung genutzt werden, die bereits Filteralgorithmen, die ursprünglich im Kontext hyperbolischer Erhaltungsgleichungen konstruiert wurden, erfolgreich zur Bildwiederherstellung einsetzt, sondern es kann auch der umgekehrte Weg eingeschlagen werden. Mit Hilfe von aus der Bildverarbeitung stammenden Filterklassen lassen sich, wie hier gezeigt, beispielsweise TV-Eigenschaften von Filtern, entwickelt im Kontext hyperbolischer Erhaltungsgleichungen, erklären und verändern. Darüber hinaus eröffnen bekannte und bewährte Filterklassen der Bildverarbeitung ein weites Spektrum neuer Filtermethoden für das numerischen Lösen hyperbolischer Erhaltungsgleichungen.

In den folgenden Kapiteln untersuchen wir verschiedene Filteralgorithmen der Bildrekonstruktion im Hinblick auf ihre Anwendung auf Erhaltungsgleichungen und entwickeln, aufbauend auf Ideen der Bildverarbeitung, neue nichtlineare Filtermethoden.



# 3 Nichtlineare, diskrete Filter der Bildverarbeitung

Im vorangehenden Kapitel haben wir beispielhaft an den ersten diskreten Filteralgorithmen von Engquist, Lötstedt und Sjögreen gezeigt, daß ihre kontinuierlichen Darstellungen vergleichbar mit Filterklassen der Bildverarbeitung sind. Die Parallelen zwischen beiden Themengebieten erlauben es uns, Informationen über Filtereigenschaften aus dem Kontext der Bildverarbeitung auf Filtermethoden für hyperbolische Erhaltungsgleichungen zu übertragen und so zu verbessern.

Eine andere Möglichkeit, Erkenntnisse der Bildverarbeitung auf dem Gebiet der Erhaltungsgleichungen zu nutzen, ist die Übertragung und Anpassung digitaler Filter aus der Signalverarbeitung auf unsere Problemstellung.

Im Bereich der numerischen Methoden für Erhaltungsgleichungen spielt der Begriff der Totalvariation eine übergeordnete Rolle. Sie ist ein Maß für die Oszillationen numerischer Lösungen, die es zu verkleinern gilt. Dies gewährleisten die sogenannten TV-Filter, die die Totalvariation eines Bildes beim Filtervorgang verkleinern oder minimieren. Da wir dieses Filterverhalten auch beim numerischen Lösen hyperbolischer Erhaltungsgleichungen anstreben, wählen wir aus den Algorithmen der Bildverarbeitung gerade die Klasse der TV-Filter aus und untersuchen ihre Anwendbarkeit im Bereich der Erhaltungsgleichungen.

Das vorliegende Kapitel unterteilt sich in vier Abschnitte, wobei die ersten beiden der Übertragung bekannter TV-Filter der Bildverarbeitung auf den Kontext hyperbolischer Erhaltungsgleichungen gewidmet sind. Die Implementierung der resultierenden Lösungsverfahren in ein- und zweidimensionale Testfälle demonstriert die Leistungsfähigkeit dieser Algorithmen, zeigt aber auch, daß einige der Verfahren Verschiebungsfehler in den Lösungen produzieren, die auf die jeweilige Diskretisierung zurückzuführen sind. Daher entwickeln wir im dritten Abschnitt neue Diskretisierungen für die zugrundeliegende Modellgleichung und untersuchen ihre numerischen Eigenschaften. Da jedes resultierende Verfahren eine diskrete Version der Gleichung  $u_t = (\operatorname{sgn} u_x)_x$  enthält, analysieren wir die mathematischen und numerischen Eigenschaften dieser Gleichung im letzten Abschnitt des Kapitels.

### 3.1 Digitaler TV–Filter von Chan, Osher und Shen

Die Theorie der TV–Filter wird erstmalig in der Doktorarbeit von Rudin [56] dargestellt und später von Osher und seinen Mitarbeitern weiter entwickelt. Als ein Ergebnis dieser Arbeiten veröffentlichten Chan, Osher und Shen 1999 [12] einen digitalen Finite–Differenzen–Filter auf Graphen, der die Fähigkeit besitzt, digitale Bilder zu entrauschen, ohne dabei Kanten oder andere Bildstrukturen zu verwischen.

Diesen Filteralgorithmus wollen wir im Folgenden beschreiben und zeigen, wie er im Kontext hyperbolischer Erhaltungsgleichungen als einfache, aber wirkungsvolle Dissipation interpretiert und implementiert werden kann.

#### 3.1.1 Das klassische *TV restoration model* von Rudin, Osher und Fatemi

Der digitale TV–Filter von Chan, Osher und Shen beruht auf dem *TV restoration model* von Rudin, Osher und Fatemi, einem der erfolgreichsten Werkzeuge zur Bildwiederherstellung und Strukturverstärkung in der Bildverarbeitung [53], [55]. Zum besseren Verständnis der Filterkonstruktion wird dieses Modell, das auf dem, in Kapitel 2.1.1 dargestellten, Variationsansatz beruht, kurz hergeleitet und erläutert. Dabei interpretieren wir die Modellvorstellung gleich im Kontext numerischer Lösungen hyperbolischer Erhaltungsgleichungen.

Wir betrachten eine numerische Lösung  $\tilde{u}$  zu einem festen Zeitpunkt  $t$  als Funktion von  $x$  und  $y$ , also  $\vec{x} := (x, y) \mapsto \tilde{u}(x, y)$ ,  $\vec{x} \in \Omega$ , wobei  $\Omega$  das betrachtete Rechengebiet bezeichnet. Da wir zur Approximation einfache Differenzenverfahren höherer Ordnung verwenden, ist die numerische Lösung  $\tilde{u}$  mangels Dämpfung durch Oszillationen verfälscht. Daher betrachten wir die numerische Lösung als Zerlegung

$$\tilde{u}(\vec{x}) = u(\vec{x}) + n(\vec{x}),$$

wobei  $u$  die tatsächliche, physikalische Lösung darstellt und  $n$  zufälliges Rauschen mit

$$En(\vec{x}) = 0, \tag{3.1}$$

$$En^2(\vec{x}) = \sigma^2, \tag{3.2}$$

das heißt, Erwartungswert Null und Varianz  $\sigma^2$ , also Standardabweichung  $\sigma$ , bezeichnet. Dabei ist zu beachten, daß numerische Lösungen von Erhaltungsgleichungen, im Gegensatz zu verrauschten Bildern der Bildverarbeitung, die Bedingung (3.2) nicht global erfüllen, da Oszillationen nur in der Nähe von Unstetigkeiten auftreten, nicht in glatten Lösungsregionen. Um in der numerischen Lösung  $\tilde{u}$  die unphysikalischen Oszillationen zu glätten und die tatsächliche Lösung  $u$  wiederherzustellen, ist es unser Ziel, die Totalvariation

$$TV(u) := \int_{\Omega} |\nabla u| \, d\vec{x}$$

unter den Nebenbedingungen (3.1) und (3.2) zu minimieren. Diese Bedingungen an  $n$  lassen sich umformulieren zu

$$\int_{\Omega} u(\vec{x}) d\vec{x} = \int_{\Omega} \tilde{u}(\vec{x}) d\vec{x}, \quad (3.3)$$

$$\frac{1}{2} \int_{\Omega} (u - \tilde{u})^2 d\vec{x} = \sigma^2, \quad (3.4)$$

wobei das Einfügen des Faktors  $\frac{1}{2}$  im Folgenden klar wird.

Da wir unseren Filter konservativ in das Differenzenverfahren implementieren, wird die erste Bedingung (3.3) automatisch erfüllt und wir brauchen sie nicht weiter berücksichtigen. Um das Minimierungsproblem unter der Nebenbedingung (3.4) zu lösen, führen wir einen Lagrangeschen Multiplikator  $\lambda$  ein, so daß

$$J[u] := \int_{\Omega} |\nabla u| d\vec{x} + \frac{\lambda}{2} \int_{\Omega} (u - \tilde{u})^2 d\vec{x}$$

zu minimieren ist.  $J[u]$  läßt sich schreiben als

$$J[u] = \int_{\Omega} F(u, \partial_x u, \partial_y u) d\vec{x}$$

mit

$$F(u, \partial_x u, \partial_y u) := |\nabla u| + \frac{\lambda}{2} (u - \tilde{u})^2.$$

Als Ansatz wählen wir

$$u(x, y, \xi) := u_0(x, y) + \xi h(x, y),$$

wobei  $h$  eine Funktion aus einem Raum  $\mathbb{H}$  von zulässigen Variationen ist, das heißt, wir haben speziell  $h|_{\partial\Omega} = 0$ . Die notwendige Bedingung dafür, daß  $u_0$  ein relatives Minimum von  $J$  ist, lautet

$$\forall h \in \mathbb{H} : DJ[u](h) = 0.$$

Wir schreiben  $J[u]$  zu

$$\bar{J}(\xi) = \int_{\Omega} F(u_0 + \xi h, \partial_x u_0 + \xi \partial_x h, \partial_y u_0 + \xi \partial_y h) dx$$

um und berechnen das Gateaux-Differential

$$\frac{d\bar{J}}{d\xi}(\xi) = \int_{\Omega} \partial_1 F \cdot h + \partial_2 F \cdot \partial_x h + \partial_3 F \cdot \partial_y h dx,$$

wobei zur Übersichtlichkeit jeweils die Argumente weggelassen wurden. Da  $DJ[u_0](h) = \frac{d\bar{J}}{d\xi}|_{\xi=0}$  gilt, liefert die notwendige Bedingung

$$\begin{aligned} \int_{\Omega} \{ \partial_1 F(u_0, \partial_x u_0, \partial_y u_0) \cdot h + \partial_2 F(u_0, \partial_x u_0, \partial_y u_0) \cdot \partial_x h \\ + \partial_3 F(u_0, \partial_x u_0, \partial_y u_0) \cdot \partial_y h \} dx = 0. \end{aligned} \quad (3.5)$$

Da

$$\begin{aligned}\partial_x(\partial_2 F(u_0, \partial_x u_0, \partial_y u_0) \cdot h(x, y)) &= h \cdot \partial_x \partial_2 F(u_0, \partial_x u_0, \partial_y u_0) \\ &\quad + \partial_2 F(u_0, \partial_x u_0, \partial_y u_0) \cdot \partial_x h, \\ \partial_y(\partial_3 F(u_0, \partial_x u_0, \partial_y u_0) \cdot h(x, y)) &= h \cdot \partial_y \partial_3 F(u_0, \partial_x u_0, \partial_y u_0) \\ &\quad + \partial_3 F(u_0, \partial_x u_0, \partial_y u_0) \cdot \partial_y h\end{aligned}$$

gilt, können wir die letzten beiden Terme in (3.5) ersetzen und erhalten

$$\int_{\Omega} \{\partial_1 F - \partial_x \partial_2 F - \partial_y \partial_3 F\} \cdot h \, d\vec{x} + \int_{\Omega} \partial_x(\partial_2 F \cdot h) + \partial_y(\partial_3 F \cdot h) \, d\vec{x} = 0.$$

Anwendung des Gaußschen Integralsatzes auf den zweiten Summanden ergibt

$$\int_{\Omega} \{\partial_1 F - \partial_x \partial_2 F - \partial_y \partial_3 F\} \cdot h \, d\vec{x} + \oint_{\partial\Omega} (\partial_2 F, \partial_3 F) \cdot \begin{pmatrix} n_x \\ n_y \end{pmatrix} \cdot h \, ds = 0,$$

wobei  $\vec{n} = (n_x, n_y)^T$  der äußere Einheitsnormalenvektor an  $\partial\Omega$  ist. Da  $h$  eine zulässige Variation ist, folgt aus  $h|_{\partial\Omega} = 0$ , daß das Linienintegral verschwindet. Der Fundamentalsatz der Variationsrechnung liefert dann die Euler–Lagrange–Gleichung

$$\partial_1 F - \partial_x \partial_2 F - \partial_y \partial_3 F = 0. \quad (3.6)$$

Für unser spezielles  $F$  erhalten wir

$$\partial_1 F = \lambda(u - \tilde{u}), \quad \partial_2 F = \frac{\partial_x u}{|\nabla u|}, \quad \partial_3 F = \frac{\partial_y u}{|\nabla u|},$$

und damit lautet die Euler–Lagrange–Gleichung unseres bedingten TV–Minimierungsproblems

$$\partial_x \left( \frac{\partial_x u}{|\nabla u|} \right) + \partial_y \left( \frac{\partial_y u}{|\nabla u|} \right) = \lambda(u - \tilde{u}),$$

oder kürzer

$$\nabla \cdot \left( \frac{\nabla u}{|\nabla u|} \right) = \lambda(u - \tilde{u}).$$

Für numerische Zwecke kann diese Gleichung mittels einer zeitlichen Hilfsvariable in die Form einer *steepest descent*–Methode umformuliert werden zu

$$\partial_t u = \nabla \cdot \left( \frac{\nabla u}{|\nabla u|} \right) + \lambda(\tilde{u} - u). \quad (3.7)$$

Im Fall  $\lambda = 0$  ist dies die berühmte TV–Diffusion von Osher und Rudin, beschrieben und untersucht beispielsweise in [43] und [76].

Rudin, Osher und Fatemi haben dieses Modell (3.7) bereits auf kartesischen Gittern mit Hilfe klassischer finiter Differenzen diskretisiert und beachtenswerte Resultate beim Entrauschen von Bildern erzielt.

Bevor wir im Folgenden den diskreten Filter und seine numerische Implementierung für unsere Anwendungen formulieren, sei darauf hingewiesen, daß sich bei der Diskretisierung der Gleichung (3.7) Probleme ergeben, wenn der Ausdruck  $|\nabla u|$  Null wird, was für näherungsweise konstantes  $u$  der Fall ist. Um diese Probleme zu umgehen, wird für numerische Berechnungen eine regularisierte Version des Gradienten,

$$|\nabla u|_{\varepsilon} := \sqrt{|\nabla u|^2 + \varepsilon^2},$$

eingeführt, wobei  $\varepsilon > 0$  ein betragsmäßig kleiner Parameter sei.

### 3.1.2 Filterbeschreibung

Chan, Osher und Shen betrachten einen ungerichteten Graphen mit endlicher Knotenmenge  $V$  und Kantenmenge  $E$ , den wir als Gitter, bestehend aus Knoten und Kanten, auffassen können. Jede Kante  $e \in E$  verbindet dabei zwei Knoten  $i, j \in V$ , die wir als Gitterpunkte auffassen. Wir bezeichnen mit

$$\sum_{e \vdash i}$$

die Summe über alle Kanten  $e$ , die vom Knoten  $i$  ausgehen und mit

$$\sum_{j \sim i}$$

die Summe über alle Knoten  $j$ , die mit dem Knoten  $i$  durch eine Kante verbunden sind. Die Abbildung  $u : V \rightarrow \mathbb{R}$  sei ein digitales Signal und  $u_i$  bezeichne den Signalwert am Knoten  $i$ . Die *diskrete Lokalvariation* am Knoten  $i$  wird durch

$$|\nabla_i u| = \sqrt{\sum_{j \sim i} (u_j - u_i)^2}$$

erklärt, ihre *regularisierte Variation* durch

$$|\nabla_i u|_\varepsilon = \sqrt{|\nabla_i u|^2 + \varepsilon^2}.$$

Führen wir die *Kantenableitung* entlang der Kante  $j \sim i$  gemäß

$$\partial_e u|_i := u_j - u_i$$

ein, läßt sich die diskrete Lokalvariation darstellen durch

$$|\nabla_i u| = \sqrt{\sum_{e \vdash i} (\partial_e u|_i)^2},$$

wobei entlang der Kante  $j \sim i$  die Gleichung  $\partial_e u|_i = -\partial_e u|_j$  gilt.

Der nichtlineare diskrete Filter von Chan, Osher und Shen verwendet für die Berechnung eines neuen, gefilterten Wertes von  $u_i$  die Werte  $u_j$  in der Nachbarschaft von  $u_i$  sowie den verrauschten Ausgangswert  $\tilde{u}_i$ . Der zugrundeliegende, zu entrauschende Datensatz geht daher bei jedem Filterschritt des nichtlinearen Filters wieder in die Berechnung ein, was eine wesentliche Eigenschaft dieses Filtermodells ist. Der Filter hängt von zwei einzustellenden Parametern ab, einem Regularisierungsparameter  $\varepsilon > 0$  und einem sogenannten Anpassungsparameter  $\lambda$ , auf deren Wahl wir später genauer eingehen werden.

Für ein Anfangssignal  $\tilde{u}$ , das die verrauschte Version des tatsächlichen Signals darstellt, ist der datenabhängige digitale TV-Filter  $\mathcal{F}^{\lambda, \varepsilon} : u \mapsto u^{\text{neu}}$  für jeden Knoten  $i$  gegeben durch

$$u_i^{\text{neu}} = \mathcal{F}_i^{\lambda, \varepsilon}(u) := \sum_{j \sim i} h_{ij}(u) u_j + h_{ii}(u) \tilde{u}_i, \quad (3.8)$$

wobei  $u$  das jeweils existierende Signal bezeichnet und  $u^{\text{neu}}$  der Rückgabewert des Filters, also das entrauschte Signal, ist.

Die Filterkoeffizienten berechnen sich gemäß

$$h_{ij}(u) = \frac{w_{ij}(u)}{\lambda + \sum_{k \sim i} w_{ik}(u)}, \quad h_{ii}(u) = \frac{\lambda}{\lambda + \sum_{k \sim i} w_{ik}(u)} \quad (3.9)$$

mit

$$w_{ij}(u) = \frac{1}{|\nabla_i u|_\varepsilon} + \frac{1}{|\nabla_j u|_\varepsilon}, \quad (3.10)$$

wobei an jedem Knoten  $i$  gilt:  $h_{ii} + \sum_{j \sim i} h_{ij} = 1$ .

Ein einzelner Filterschritt kann damit algorithmisch wie folgt beschrieben werden:

- für alle Knoten  $i$ 
  - berechne die regularisierte diskrete Lokalvariation  $|\nabla_i u|_\varepsilon$  sowie die Variationen  $|\nabla_j u|_\varepsilon$  aller Nachbarknoten  $j$  von  $i$
  - berechne die Gewichtungen  $w_{ij}$  und die Filterkoeffizienten  $h_{ii}$  und  $h_{ij}$
  - berechne  $u_i^{\text{neu}} = \mathcal{F}^{\lambda, \varepsilon}(u)$
- Ende Schleife

Dieser Filterschritt wird bei den Anwendungen der Bildverarbeitung so lange wiederholt, bis eine stationäre Lösung, deren Existenz wir voraussetzen, erreicht wird.

Der diskrete, nichtlineare Tiefpaßfilter besitzt zwei einstellbare Parameter, einen betragsmäßig kleinen, positiven Regularisierungsparameter  $\varepsilon$ , der die numerische Stabilität gewährleistet, und einen positiven Anpassungsparameter  $\lambda$ . Während  $\varepsilon$  lediglich eine technisch begründete Konstante ist, ist der Anpassungsparameter  $\lambda$  ausschlaggebend für die Wiederherstellung verrauschter Kanten und anderer Bildstrukturen. Der Wert von  $\lambda$  ist vergleichbar mit  $1/\sigma^2$ , wobei  $\sigma^2$  die Varianz des zugrundeliegenden Rauschens ist. Damit fungiert  $\lambda$  als Indikator für Rauschen, das geglättet werden soll, oder wirkliche Intensitätssprünge, die nicht durch den Filter zerstört werden dürfen.

Vor diesem Hintergrund läßt sich die Adaptivität des digitalen Filters leicht verstehen. Eine starke lokale Variation  $|\nabla u|$ , groß genug, um vom Rauschen unterschieden zu werden, weist auf einen Sprung oder eine Kante, übernommen von  $\tilde{u}$ , hin und soll dementsprechend erhalten bleiben. Der Filter gewährleistet dieses Verhalten, da ein großes  $|\nabla u|$  zu einem kleinen  $w_{ij}$  (verglichen mit  $\lambda$ ) führt, so daß die Filterkoeffizienten  $h_{ii}$  fast eins und  $h_{ij}$  fast Null werden. Dies bewirkt, wie erwartet, die Sprungerhaltung. Der Fall einer kleinen lokalen Variation  $|\nabla u|$  weist auf eine flache Region hin, die in  $\tilde{u}$  verrauscht dargestellt ist und mit Hilfe des Filters entrauscht werden soll. Tatsächlich liefert ein kleines  $|\nabla u|$  ein großes  $w_{ij}$  (verglichen mit  $\lambda$ ), so daß  $h_{ii}$  fast Null und  $h_{ij}$  fast eins wird, was eine Glättung der Daten bewirkt.

Für die Bildentrauschung und –wiederherstellung hat sich gezeigt, daß in der Praxis 60 bis 80 Filteriterationen notwendig sind, um gute numerische Ergebnisse zu erzielen.

### 3.1.3 Zusammenhang zwischen diskretem Filter und kontinuierlicher TV–Diffusionsgleichung

Im folgenden Abschnitt wollen wir den Zusammenhang zwischen dem diskreten Filteralgorithmus (3.8) und dem kontinuierlichen TV–Diffusionsmodell (3.7) untersuchen.



Wie im vorangegangenen Kapitel beschrieben, wird ein einzelner Filterschritt des nichtlinearen TV-Filters bei den Anwendungen der Bildverarbeitung so oft wiederholt, bis eine stationäre Lösung erreicht wird, deren Existenz wir voraussetzen. Aufgrund dieser Annahme existiert ein Fixpunkt  $u_i = \mathcal{F}^{\lambda, \varepsilon}(u)$  der Iteration  $u_i^{\text{neu}} = \mathcal{F}^{\lambda, \varepsilon}(u)$ , und mit der diskreten Filtergleichung (3.8) ergibt sich

$$u_i = \sum_{j \sim i} h_{ij}(u) u_j + h_{ii}(u) \tilde{u}_i = \sum_{j \sim i} \frac{w_{ij}(u)}{\lambda + \sum_{k \sim i} w_{ik}(u)} u_j + \frac{\lambda}{\lambda + \sum_{k \sim i} w_{ik}(u)} \tilde{u}_i.$$

Umformulierungen ergeben

$$\left( \lambda + \sum_{k \sim i} w_{ik}(u) \right) u_i = \sum_{j \sim i} w_{ij}(u) u_j + \lambda \tilde{u}_i$$

oder

$$\left( \lambda + \sum_{j \sim i} w_{ij}(u) \right) u_i = \sum_{j \sim i} w_{ij}(u) u_j + \lambda \tilde{u}_i,$$

und Termumordnung liefert

$$\sum_{j \sim i} w_{ij}(u_j - u_i) + \lambda(\tilde{u}_i - u_i) = 0. \quad (3.11)$$

Um die Verbindung zwischen dem diskreten Filter und dem kontinuierlichen Modell (3.7) herzustellen, betrachten wir die stationäre Gleichung (3.11) als diskrete Differenzengleichung auf dem, durch die Knoten und Kanten definierten, zugrundeliegenden Graphen.

**Lemma 3.1**

$$\sum_{j \sim i} w_{ij}(u_j - u_i) = \sum_{e \vdash i} \partial_e \left( \frac{-1}{|\nabla u|_\varepsilon} \partial_e u \right).$$

**Beweis** Die linke Seite der Gleichung läßt sich mittels der Definition der Gewichtungen (3.10) umschreiben zu

$$\sum_{j \sim i} \left( \frac{1}{|\nabla_i u|_\varepsilon} + \frac{1}{|\nabla_j u|_\varepsilon} \right) (u_j - u_i).$$

Die rechte Seite kann mit Hilfe der Definition der diskreten Kantenableitung umgeformt werden zu

$$\begin{aligned} \sum_{e \vdash i} \partial_e \left( \frac{-1}{|\nabla u|_\varepsilon} \partial_e u \right) &= \sum_{j \sim i} \frac{-1}{|\nabla_j u|_\varepsilon} \partial_e u|_j - \frac{-1}{|\nabla_i u|_\varepsilon} \partial_e u|_i \\ &= \sum_{j \sim i} \frac{-1}{|\nabla_j u|_\varepsilon} (u_i - u_j) + \frac{1}{|\nabla_i u|_\varepsilon} (u_j - u_i) \\ &= \sum_{j \sim i} \left( \frac{1}{|\nabla_i u|_\varepsilon} + \frac{1}{|\nabla_j u|_\varepsilon} \right) (u_j - u_i), \end{aligned}$$

und dies ist gerade die linke Seite. ■

Aufgrund des Lemmas ist der digitale TV-Filter von Chan, Osher und Shen für den stationären Fall durch die finite Differenzengleichung

$$\sum_{e \vdash i} \partial_e \left( \frac{-1}{|\nabla u|_\varepsilon} \partial_e u \right) + \lambda(\tilde{u}_i - u_i) = 0 \quad (3.12)$$

auf dem Graphen  $\{e \mid e \vdash i\}$  gegeben. Hier läßt sich nun der Zusammenhang zwischen dem diskreten Filter (3.8) und dem kontinuierlichen Modell (3.7) erkennen. Die diskrete Differenzengleichung (3.12), betrachtet auf einem Graphen, ist die diskrete Version der kontinuierlichen TV-Diffusionsgleichung (3.7).

### 3.1.4 Numerische Ergebnisse für die Interpretation im Kontext hyperbolischer Erhaltungsgleichungen

Für die Anwendung des TV-Filters im Kontext numerischer Lösungen für Erhaltungsgleichungen interpretieren wir das verrauschte Signal  $\tilde{u}$  als eine oszillierende Lösung einer hyperbolischen Erhaltungsgleichung. Im Gegensatz zu der in der Bildverarbeitung gegebenen Situation haben wir nicht mehr globales zufälliges Rauschen auf unseren Daten, sondern lokale Oszillationen in der Umgebung von Sprüngen oder Stößen. Um Fehler zu vermeiden, die aus einer globalen Varianzannahme resultieren würden, haben wir verschiedene Routinen entwickelt, die datenabhängig eine lokale Varianz und damit ein lokales  $\lambda$  für jeden Gitterpunkt berechnen. Unsere numerischen Experimente zeigen allerdings, daß nicht in jeder Situation ein lokales  $\lambda$  einem globalen vorzuziehen ist. Darüber hinaus zeigen unsere Anwendungen, daß wir schon mit wesentlich weniger als 60 bis 80 Filteriterationen sehr gute numerische Resultate erzielen können. Beispielsweise benötigen wir nach Erhalt einer oszillierenden stationären Lösung, berechnet mit einem einfachen Differenzenverfahren höherer Ordnung, lediglich eine Filteriteration, um wesentlich verbesserte Lösungen zu erhalten.

Im Folgenden gehen wir genauer auf verschiedene Anwendungen des digitalen TV-Filters im Kontext hyperbolischer Erhaltungsgleichungen ein.

Zunächst betrachten wir als Testfall das zweidimensionale Randwertproblem (1.61), dessen exakte Lösung bereits in Kapitel 1.3.2 hergeleitet wurde und in Abbildung 1.4 dargestellt ist. Die exakte Lösung dient uns dabei als Vergleichsmöglichkeit, um die Qualität der numerischen Lösungen des Randwertproblems zu beurteilen.

Wir berechnen die approximative Lösung von (1.61) mit Hilfe des oszillierenden Lax-Wendroff-Verfahrens, das wir im Rahmen einer Finite-Volumen-Approximation auf dem Sekundärgitter einer Triangulierung implementieren. Die numerische Flußfunktion des Lax-Wendroff-Verfahrens ist dabei in der Form

$$F = \frac{1}{2} (f(U_L) + f(U_R)) n_1 + \frac{1}{2} (g(U_L) + g(U_R)) n_2 - \frac{1}{4} \lambda (A n_1 + B n_2)^2 (U_R - U_L) \quad (3.13)$$

gegeben, wobei  $U_L$  und  $U_R$  die jeweils linken und rechten Zustände bezeichnen.  $A$  und  $B$  sind die an der Stelle  $\frac{1}{2}(U_L + U_R)$  berechneten Jacobi-Matrizen der Flußfunktionen  $f$  und  $g$ , wobei

für das skalare Testproblem (1.61)  $A(U) = U$  und  $B(U) = 1$  gilt. Der Vektor  $n = (n_1, n_2)^T$  bezeichnet die äußere Einheitsnormale auf dem Rand des betrachteten Kontrollvolumens, und der Faktor  $\nu = k/h$  ist durch die Bedingung

$$\text{CFL} = \nu \max_{W \in \{U_L, U_R, (U_L+U_R)/2\}} |W n_1 + n_2| \leq 1$$

gegeben. Als Rechengitter verwenden wir eine Triangulierung aus 3002 Punkten beziehungsweise 5802 Dreiecken, und als CFL-Zahl wählen wir  $\text{CFL} = 0.9$ .

Weitere Ausführungen zu Finite-Volumen-Verfahren auf unstrukturierten Gittern sind beispielsweise in [17] und [61] zu finden.

Die resultierende stationäre Lösung ist als Draufsicht mit Hilfe von Isolinen in Abbildung 3.1, in der auch die verwendete Triangulierung zu erkennen ist, dargestellt. Rechts und links vom Stoß lassen sich die typischen Oszillationen, die das Lax-Wendroff-Verfahren verursacht, identifizieren.

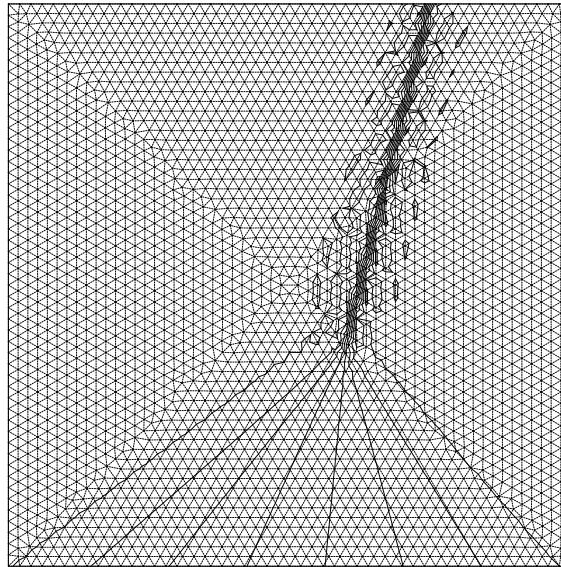


Abbildung 3.1: Oszillierende Lax-Wendroff-Lösung

Die zusätzliche Implementierung von 60 Iterationen des digitalen TV-Filters nach Beendigung des Lax-Wendroff-Verfahrens, zunächst mit einem global gewählten  $\lambda = 7$  und dem Regularisierungsparameter  $\varepsilon = 0.0001$ , ergibt die in Abbildung 3.2 dargestellte Lösung. Darüber hinaus zeigt Abbildung 3.3 ein Profil der ungefilterten und der global gefilterten Lösung, wobei der Schnitt durch die Punkte  $(0; 0.5)$  und  $(1; 0.5)$  verläuft. Beide Darstellungsmethoden verdeutlichen, daß die Oszillationen in der Nähe des Stoßes durch den zusätzlich implementierten TV-Filter vollständig beseitigt sind. Numerische Simulationen zeigen, daß wir das gleiche Resultat schon mit nur 10 Filteriterationen nach Beendigung des Differenzenverfahrens erzielen.

Allerdings müssen wir feststellen, daß nun auch tatsächliche Lösungsstrukturen überglättet worden sind, wie die gekrümmten Isolinen am unteren Bildrand in Abbildung 3.2 zeigen. Abhilfe schafft ein lokal berechnetes  $\lambda$ . Dazu berechnen wir für jeden Gitterpunkt der Triangulierung aus den Gitterwerten der direkten Nachbarpunkte die lokale Variation und damit

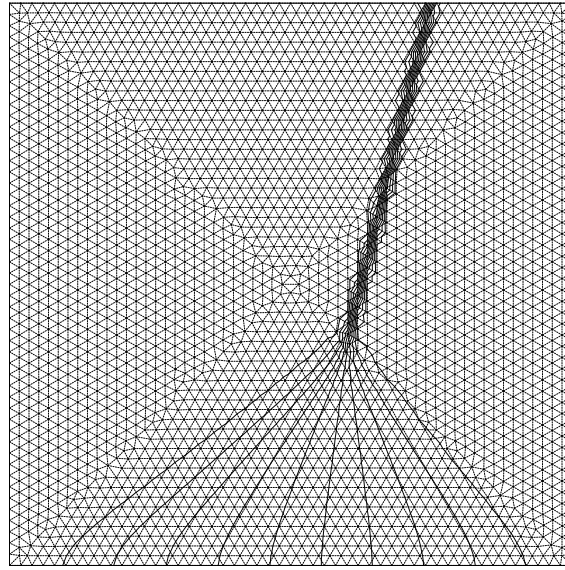
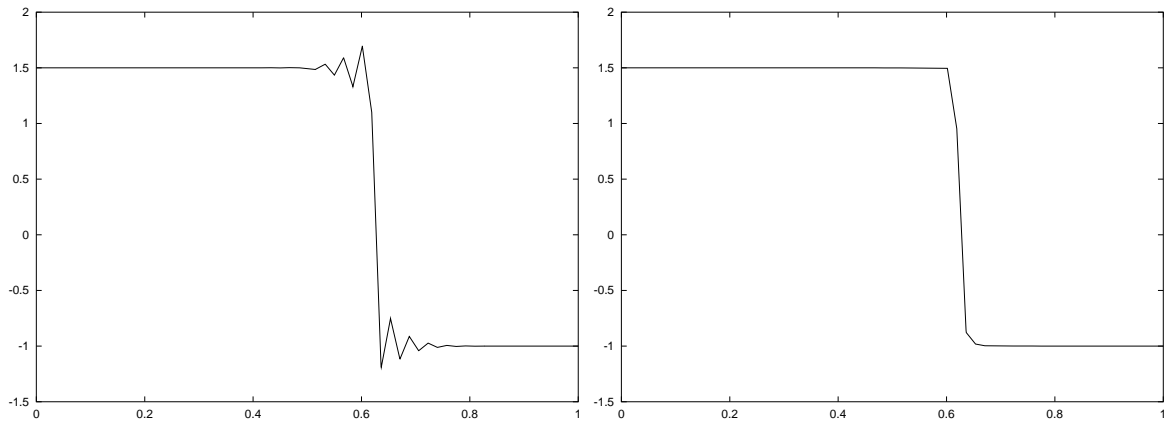


Abbildung 3.2: Global gefilterte Lax–Wendroff–Lösung

Abbildung 3.3: Schnittbilder der Lax–Wendroff–Lösung (links) und der zugehörigen global gefilterten Lösung (rechts) für einen Schnitt durch die Punkte  $(0; 0.5)$  und  $(1; 0.5)$ 

ein lokales  $\lambda$ . Für die Punkte, die genau auf dem Stoß liegen, scheitert diese Methode allerdings, da ihre Nachbarpunkte auf beiden Seiten des Stoßes liegen, und damit die tatsächliche Sprunghöhe in die Varianzberechnung, die nur das Rauschen widerspiegeln soll, mit einbezogen wird und sie daher verfälscht. Dieses Problem können wir umgehen, wenn wir zuvor einen einfachen Kantendetektor auf unseren Datensatz anwenden, der alle Punkte, die auf Kanten liegen, markiert. Bei der lokalen Varianzberechnung bleiben dann alle markierten Nachbarpunkte unberücksichtigt. Ist der Punkt, für den die Varianz berechnet werden soll, selbst markiert, so werden nur die Nachbarpunkte auf einer Seite des Stoßes betrachtet. Auf diese Weise können wir verhindern, daß Varianzberechnungen über Stöße hinweg durchgeführt werden und vermeiden so falsche lokale Werte für den Anpassungsparameter  $\lambda$ .

Abbildung 3.4 zeigt die gefilterte Lösung, bei der solch eine datenabhängige, lokale Wahl des Anpassungsparameters  $\lambda$  mit einem einfachen Laplace–Detektor zur Kantenmarkierung verwendet wurde. Wie im vorherigen Fall gilt  $\varepsilon = 0.0001$ , und es wurden 60 Filteriteratio-

nen auf die stationäre Lax–Wendroff–Lösung angewendet, wobei auch hier mit lediglich 10 Iterationen genau das gleiche Resultat erzielt werden kann. Das Schnittbild für einen Schnitt durch die Punkte  $(0; 0.5)$  und  $(1; 0.5)$  ist in Abbildung 3.5 dargestellt.

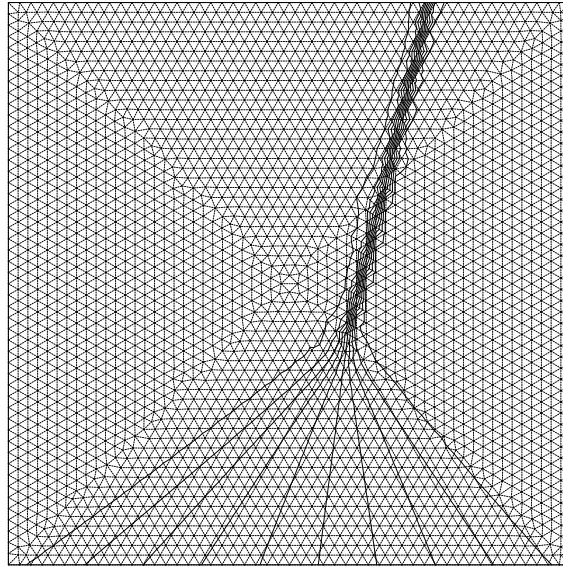


Abbildung 3.4: Lokal gefilterte Lax–Wendroff–Lösung

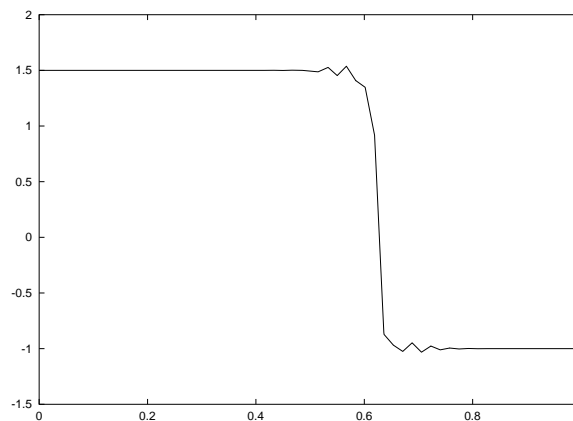


Abbildung 3.5: Schnittbild der lokal gefilterten Lax–Wendroff–Lösung für einen Schnitt durch die Punkte  $(0; 0.5)$  und  $(1; 0.5)$

Ein Vergleich mit dem ungefilterten Lax–Wendroff–Verfahren zeigt, daß die Oszillationen in der Umgebung der Unstetigkeit reduziert sind und der lineare Teil der Lösung, im Gegensatz zur globalen Filterroutine, erhalten ist. Die Restoszillationen, die in Abbildung 3.5 zu erkennen sind, lassen sich allerdings nicht weiter durch Erhöhung der Iterationsschritte verringern. Dennoch wird deutlich, daß der digitale TV-Filter nicht nur im Kontext der Bildrekonstruktion sinnvolle Anwendung findet, sondern daß er auch bei der numerischen Approximation hyperbolischer Erhaltungsgleichungen zur Glättung von Oszillationen unter Beibehaltung der Unstetigkeiten eingesetzt werden kann.

Als Testfall für Systeme von Erhaltungsgleichungen simulieren wir die Umströmung des NACA0060 Profils mit Hilfe der Eulergleichungen der Gasdynamik. Dieses Testproblem ist am Ende von Kapitel 1.3.2 beschrieben, die konservative Formulierung der Eulergleichungen ist dort unter (1.62) gegeben.

Zunächst berechnen wir die numerische Lösung dieses Testfalls mit dem Lax–Wendroff–Verfahren, das wir wieder im Rahmen einer Finite–Volumen–Approximation auf dem Sekundärgitter der Triangulierung implementieren. Die numerische Flußfunktion des Lax–Wendroff–Verfahrens ist dabei in der Form (3.13) gegeben, wobei  $A$  und  $B$  jetzt die Jacobi–Matrizen der Flußfunktionen  $f$  und  $g$ , die in Kapitel 1.3.2 gegeben sind, bezeichnen. Die CFL–Zahl setzen wir für den vorliegenden Testfall auf 0.8.

Ebenso wie im skalaren Fall liefert das Lax–Wendroff–Verfahren eine stationäre Lösung, deren Dichteverlauf in Abbildung 3.6 links unter Verwendung von Isolinien als Ansicht von oben dargestellt ist. Der linke Teil der Abbildung 3.7 zeigt darüber hinaus einen Querschnitt der Lösung oberhalb des Tragflächenprofils. Oberhalb und unterhalb des Profils haben sich zwei Stöße ausgebildet, in deren Umgebung Oszillationen zu erkennen sind.

Nun wiederholen wir die Simulation und führen nach Erhalt der stationären Lösung einen einzigen Filterschritt des nichtlinearen TV–Filters aus, indem wir die Filterroutine unabhängig voneinander in jede Systemkomponente integrieren. Fast ohne zusätzlichen Rechenaufwand erhalten wir die Lösung, die in den Abbildungen 3.6 und 3.7 jeweils rechts dargestellt ist.

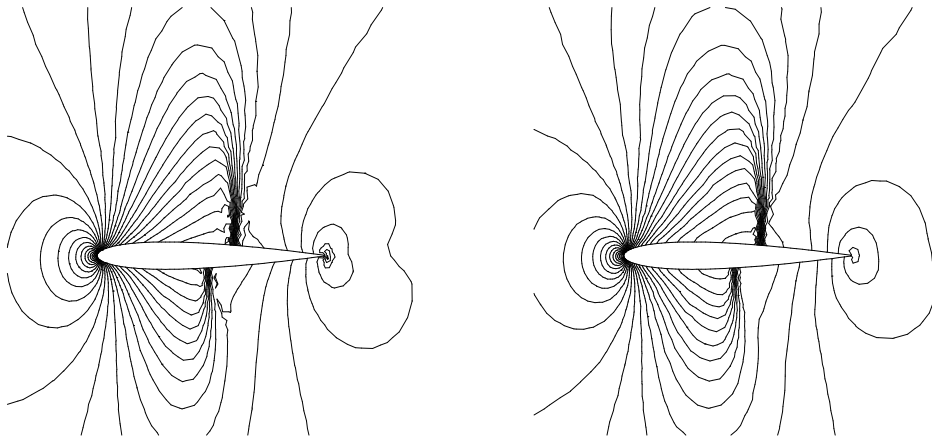


Abbildung 3.6: Lax–Wendroff–Lösung (links) und mit TV–Filter gefilterte Lax–Wendroff–Lösung (rechts)

Für die beiden einzustellenden Parameter haben wir  $\varepsilon = 0.0001$  und  $\lambda = 140$  verwendet. Der Wert für  $\lambda$ , der hier weitaus größer als in Anwendungen der Bildverarbeitung gewählt wurde, läßt sich dadurch erklären, daß wir im Gegensatz zu 60 bis 80 Filteriterationen lediglich einen einzigen Filterschritt im Anschluß an das Differenzenverfahren durchführen.

Abbildung 3.7 rechts zeigt, daß der TV–Filter auch für dieses Testproblem das Aussehen und die Steilheit von Stößen vollständig erhält und Oszillationen in ihrer Umgebung glättet. Vergleichen wir die Querschnitte durch die ungefilterte und gefilterte Lösung unterhalb des Profils, dargestellt in Abbildung 3.8, läßt sich erkennen, daß selbst die Oszillationen in der Nähe des kleineren Stoßes an der Profilunterseite durch den Filter geglättet worden sind, ohne daß der Stoß selbst zerstört wurde, was viele gängige Filteralgorithmen nicht leisten. Empirisch haben wir damit gezeigt, daß der digitale TV–Filter für das numerische Lösen hy-

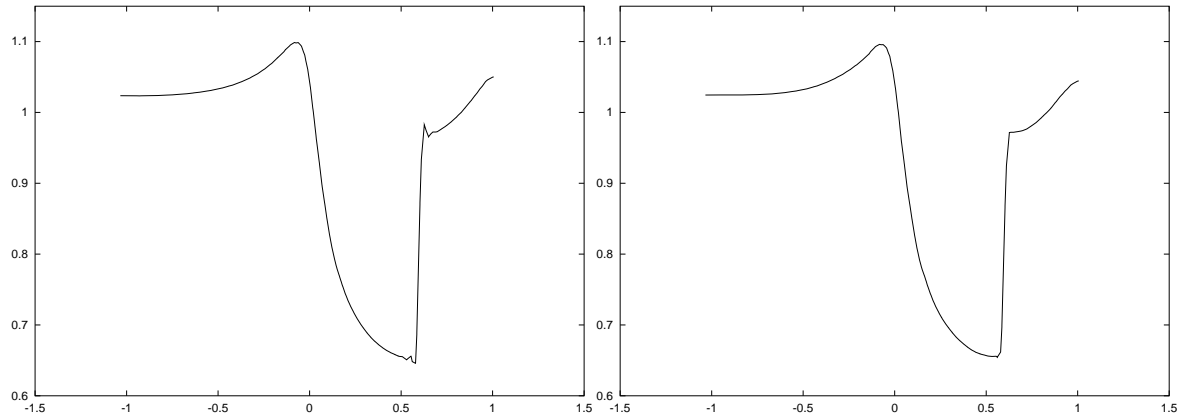


Abbildung 3.7: Schnitt durch die Lax–Wendroff–Lösung (links) und die gefilterte Lösung (rechts) oberhalb des Tragflächenprofils

perbolischer Erhaltungsgleichungen sehr gute Resultate liefert, wenn er auf stationäre Lösungen angewendet wird, die durch Oszillationen, bedingt durch Differenzenverfahren höherer Ordnung, gestört sind.

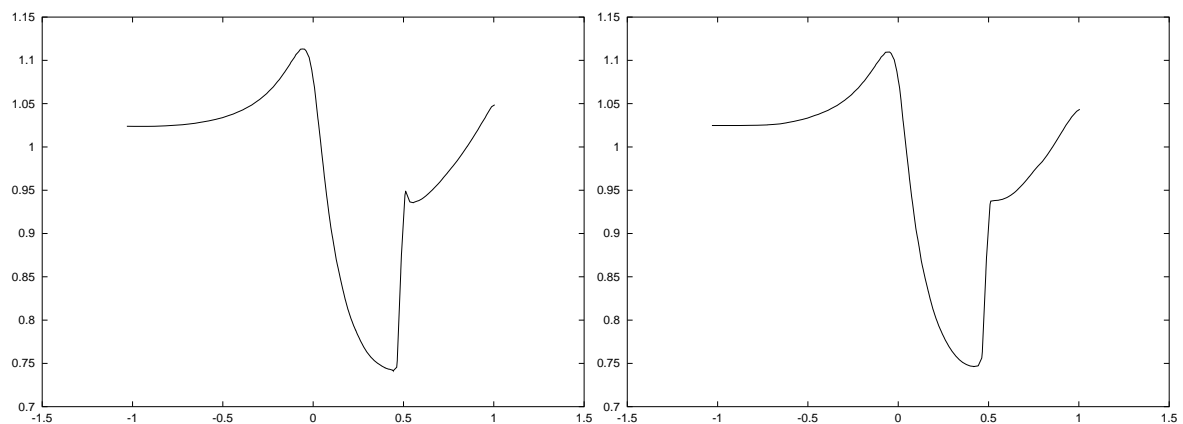


Abbildung 3.8: Schnitt durch die Lax–Wendroff–Lösung (links) und die gefilterte Lösung (rechts) unterhalb des Tragflächenprofils

Bisher haben wir den digitalen TV-Filter stets nach Beendigung des Differenzenverfahren auf die resultierende stationäre Lösung angewendet. Eine andere Möglichkeit, Filteralgorithmen in Finite-Volumen-Methoden für hyperbolische Erhaltungsgleichungen zu implementieren, ist die Anwendung der Filterroutine nach jedem Zeitschritt des Differenzenverfahrens oder jeweils nach einer festen Anzahl von Zeitschritten. Bei dieser Art der Implementierung muß das zugrundeliegende Differenzenschema nicht mehr notwendig stabil sein und eine stationäre Lösung liefern, sondern es können auch instabile Verfahren verwendet werden, die durch die regelmäßige Anwendung des Filters stabilisiert werden. Dies liefert die wünschenswerte Möglichkeit, den Rechenaufwand resultierender Verfahren weiter zu reduzieren, indem rechentechnisch aufwendige durch instabile rechenzeitgünstige Verfahren ersetzt werden. Ein typisches instabiles, aber sehr schnelles Differenzenverfahren, das an Oszillationen in Stoßnähe leidet, ist das einfache Zentrale-Differenzen-Schema zweiter Ordnung, das wir im Folgenden

für unsere Filteranwendung zugrundelegen wollen.

Wir betrachten das einfache Cauchy–Problem für Burgers’ Gleichung, das durch (1.49) in Kapitel 1 beschrieben ist und zum Zeitpunkt  $t = 1$  einen Stoß ausbildet. Dieses Cauchy–Problem lösen wir mit Hilfe der unbedingt instabilen zentralen Differenz zweiter Ordnung

$$U_i^{n+1} = U_i^n - \frac{\Delta t}{2\Delta x}(U_{i+1}^n - U_{i-1}^n)$$

auf einen Gitter aus 400 äquidistanten Gitterpunkten. Mit einer CFL–Zahl von 0.009 erhalten wir zum Zeitpunkt  $t = 2.5$  die in Abbildung 3.9 links dargestellte oszillierende Lösung des Cauchy–Problems. Wenige Zeitschritte später würde die Lösung aufgrund ihrer Oszillationen zusammenbrechen.

Fügen wir eine Filteriteration mit dem nichtlinearen TV–Filter nach jedem Zeitschritt des Differenzenverfahrens ein, so erhalten wir die im rechten Teil von Abbildung 3.9 gezeigte Lösung. Die CFL–Zahl beträgt nun 0.9 und als Parameter wurde  $\varepsilon = 0.01$  und  $\lambda = 1.0$  gewählt.

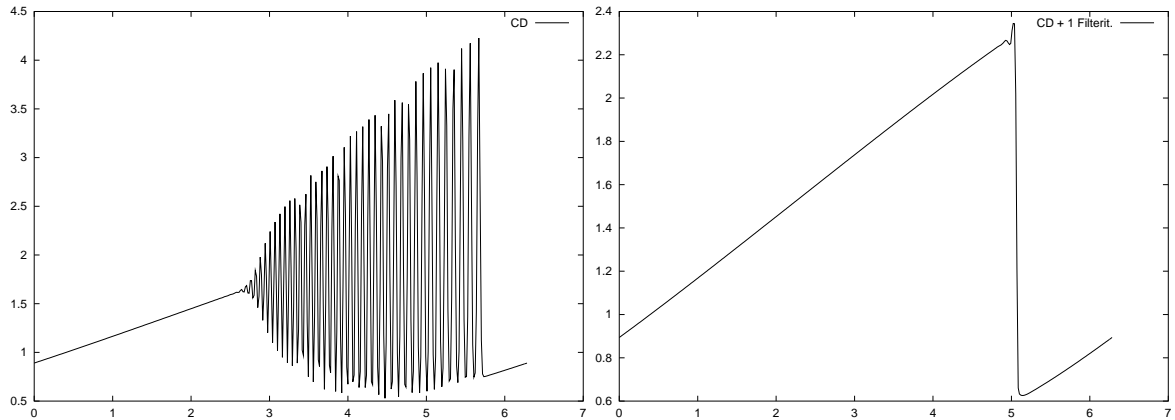


Abbildung 3.9: Ungefilterte (links) und mit einer Filteriteration pro Zeitschritt gefilterte (rechts) Zentrale–Differenzen–Lösung

Die Oszillationen, bedingt durch das Differenzenschema zweiter Ordnung, sind durch die jeweils zusätzliche Filteriteration fast vollständig beseitigt, was bei einem Vergleich mit der ungefilterten Lösung sehr beeindruckt.

Vergleichen wir die gefilterte Lösung allerdings mit einer herkömmlichen Lax–Friedrichs–Lösung mit CFL–Zahl 0.9, so stellen wir fest, daß die Lage des gefilterten Stoßes nicht mehr mit seiner tatsächlichen Lage übereinstimmt, wie Abbildung 3.10 verdeutlicht.

Erhöhen wir die Anzahl der Filteriterationen pro Zeitschritt, beobachten wir, daß der Verschiebungsfehler bei vier oder mehr Iterationen nicht mehr auftritt, wie Abbildung 3.11 zeigt. Allerdings wird die numerische Lösung gleichzeitig immer glatter, und der Filter konvergiert nicht. Die fehlende Konvergenz ist auf den zu groß gewählten Parameter  $\varepsilon = 0.01$  zurückzuführen, der bei der Berechnung lediglich das Teilen durch Null bei sehr kleinen Gradienten verhindern soll. Setzen wir  $\varepsilon = 10^{-8}$ , ergibt sich für  $\lambda = 1$  und drei Filteriterationen nach jedem Zeitschritt der zentralen Differenz das in Abbildung 3.12 links dargestellte Resultat. Jetzt konvergiert der Filter für steigende Anzahl von Filteriterationen, wie der rechte Teil der Abbildung für 100 Iterationen pro Zeitschritt zeigt. Aber auch in diesem Fall ist eine Verschiebung der Stoßlage zu beobachten, die stark von der Wahl des Parameters  $\lambda$  und der



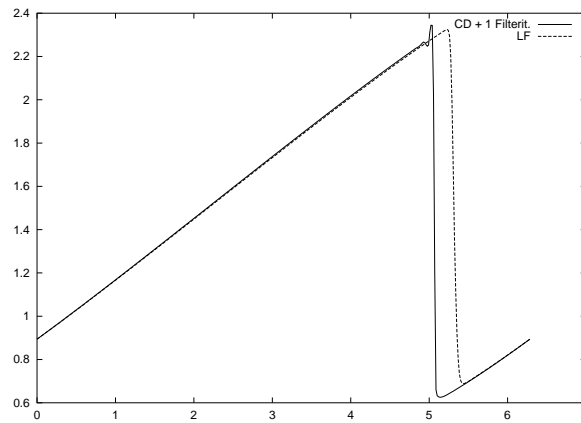


Abbildung 3.10: Gefilterte zentrale Differenz im Vergleich zur Lax-Friedrichs-Lösung

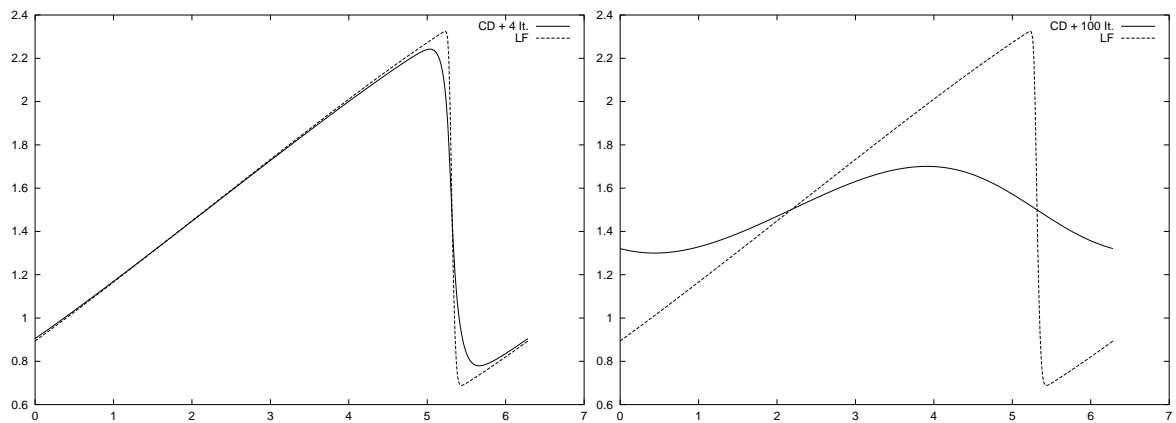
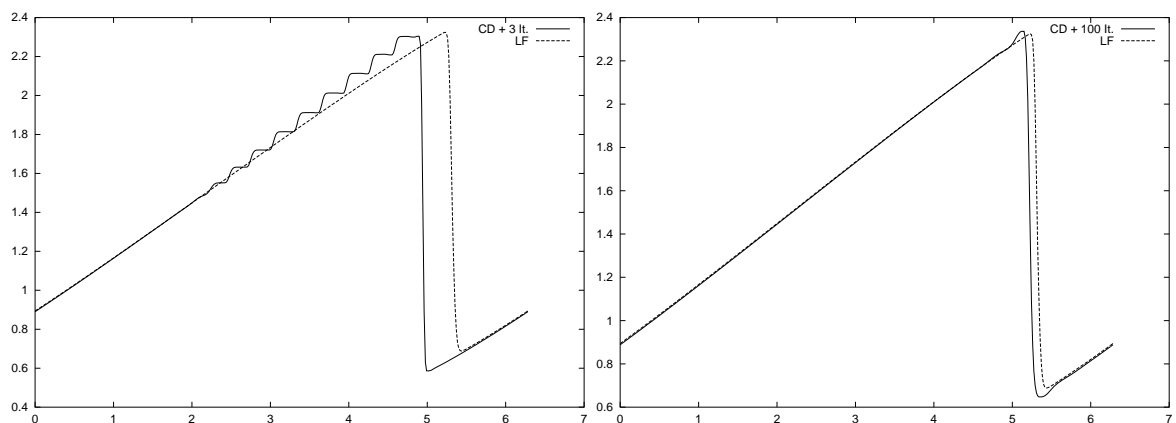


Abbildung 3.11: Gefilterte zentrale Differenz mit vier (links) bzw. 100 (rechts) Filteriterationen pro Zeitschritt im Vergleich zur Lax-Friedrichs-Lösung

Abbildung 3.12: Gefilterte zentrale Differenz mit drei (links) bzw. 100 (rechts) Filteriterationen pro Zeitschritt und  $\varepsilon = 10^{-8}$ 

Anzahl der Filteriterationen abhängt. Einerseits lassen sich Kombinationen dieser Größen

finden, für die der Verschiebungsfehler verschwindet, andererseits lassen sich die Parameter auch so einstellen, daß eine Verschiebung nach rechts statt nach links zu beobachten ist. Für das Lax–Wendroff–Verfahren als zugrundeliegende Differenzenmethode lassen sich ähnliche Eigenschaften beobachten. Bei einer CFL–Zahl von 0.9 ergibt die Implementierung einer Filteriteration nach jedem Zeitschritt mit  $\varepsilon = 10^{-8}$  und  $\lambda = 1$  das in Abbildung 3.13 links dargestellte Resultat. Ein Vergleich mit der zugehörigen Lax–Friedrichs–Lösung zeigt die

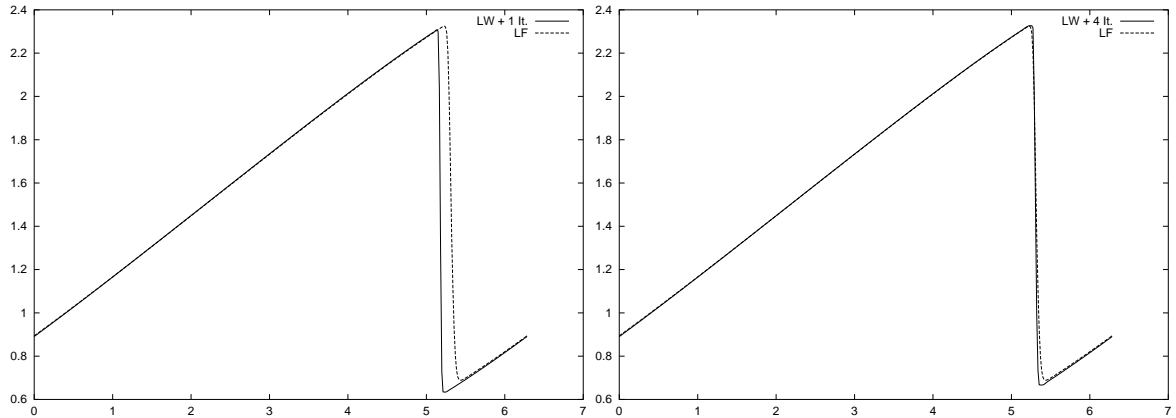


Abbildung 3.13: Vergleich zwischen der gefilterten Lax–Wendroff–Lösung mit einer (links) bzw. vier Filteriterationen pro Zeitschritt bei  $\varepsilon = 10^{-8}$  und  $\lambda = 1$  und der Lax–Friedrichs–Lösung

Verschiebung der Stoßlage, die allerdings ab vier Filteriterationen pro Zeitschritt nicht mehr auftritt, wie die rechte Darstellung in Abbildung 3.13 verdeutlicht.

Um die Ursache des Verschiebungsfehlers zu untersuchen, betrachten wir die Simulationen auf einem groben Gitter aus 40 äquidistanten Gitterpunkten und vergleichen die gefilterte Lax–Wendroff–Lösung mit der ungefilterten Lösung. Abbildung 3.14 zeigt die Ergebnisse links für eine Filteriteration und rechts für acht Iterationen pro Zeitschritt. Obwohl die kontinuierliche

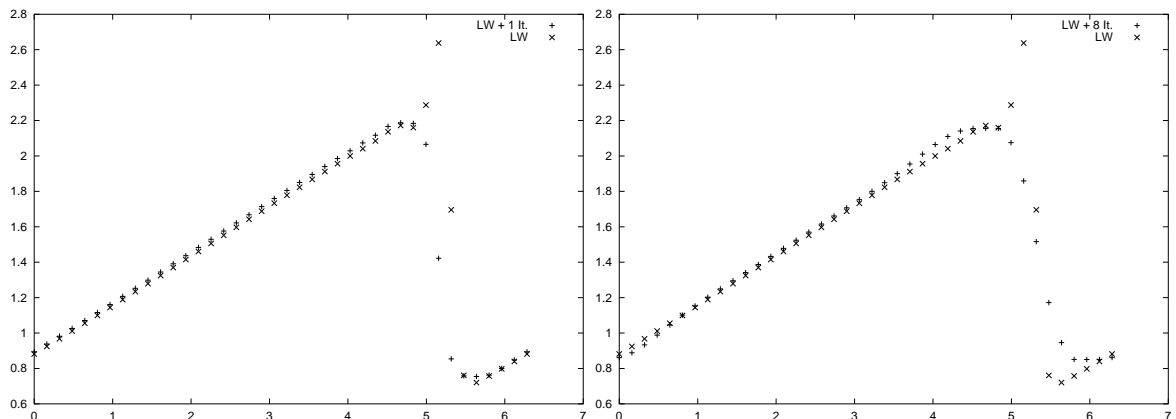


Abbildung 3.14: Gefilterte Lax–Wendroff–Lösung mit einer (links) bzw. acht Filteriterationen pro Zeitschritt im Vergleich zur ungefilterten Lax–Wendroff–Lösung

Filtergleichung

$$0 = \nabla \cdot \left( \frac{\nabla u}{|\nabla u|} \right) + \lambda(\tilde{u} - u)$$

aufgrund ihrer Divergenzform konservativ ist, und die Symmetrie der  $w_{ij}(u)$  sowie die Eigenschaft  $h_{ii} + \sum_{j \sim i} h_{ij} = 1$  vermuten läßt, daß auch der diskrete Filter (3.8) konservativ ist, zeigt die linke Darstellung von Abbildung 3.14 deutlich, daß der diskrete Filter bei Anwendung einer Filteriteration nicht konservativ ist. Die Filterroutine korrigiert die Datenpunkte im Stoß fast ausschließlich nach unten, ohne Datenpunkte an anderer Stelle zum Erhalt der Konservativität nach oben zu korrigieren. Die beobachtete Stoßverschiebung ist damit auf die Nichtkonservativität des Filters zurückzuführen. Erhöhen wir dagegen die Anzahl der Filteriterationen nach jedem Zeitschritt des Differenzenverfahrens, stellen wir fest, daß der Filter, wenn er konvergiert, konservativ ist. Die rechte Darstellung in Abbildung 3.14 deutet diese Eigenschaft schon für acht Filteriterationen pro Zeitschritt an, zum Ausgleich der nach unten korrigierten Daten im Stoß werden die Datenpunkte hinter dem Stoß nach oben korrigiert. Im Konvergenzfall ist die Konservativität exakt gegeben und die Stoßlage in der gefilterten Lax–Wendroff–Lösung stimmt mit der Stoßlage der Lax–Friedrichs–Lösung überein. Allerdings ist die Konservativität der Filtergleichung selbst bei für Konvergenz hinreichend vielen Iterationen stark von der Wahl der Filterparameter abhängig. Abbildung 3.15 zeigt beispielsweise die mit 80 Filteriterationen pro Zeitschritt gefilterte Lax–Wendroff–Lösung für  $\lambda = 20$  und  $\varepsilon = 10^{-8}$ . Die linke Darstellung zeigt, daß trotz Konvergenz des Filters nach jedem Zeitschritt

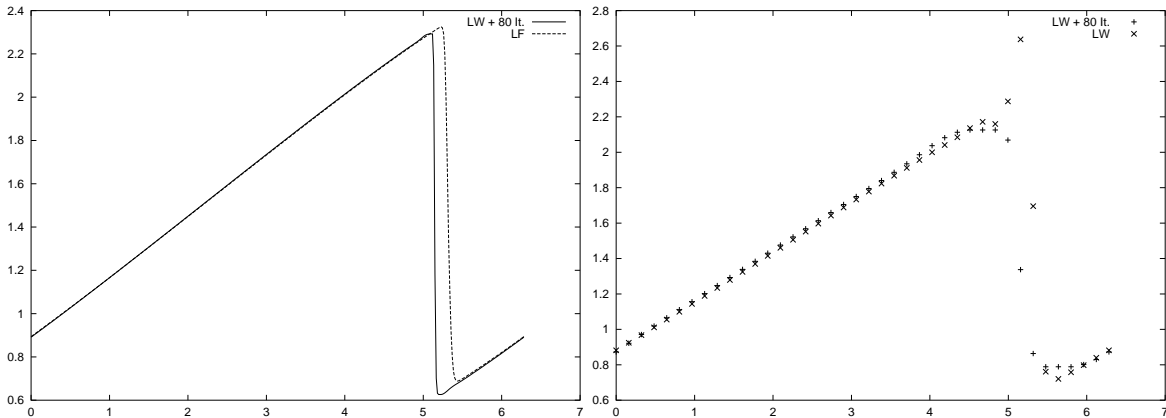


Abbildung 3.15: Mit 80 Iterationen gefilterte Lax–Wendroff–Lösung für  $\lambda = 20$  auf einem Gitter aus 400 (links) bzw. 40 Gitterpunkten (rechts) im Vergleich zur ungefilterten Lax–Wendroff–Lösung

eine Stoßverschiebung auftritt, und das Ergebnis auf dem groben Gitter verdeutlicht, daß der Filter für diese Kombination der Parameter nicht konservativ ist.

Diese Beobachtungen decken sich mit den numerischen Ergebnissen für das nach jedem Zeitschritt gefilterte Zentrale–Differenzen–Verfahren, für das wir bereits festgestellt haben, daß die Art und Größe der Stoßverschiebung abhängig von der Wahl der Filterparameter ist.

Da die Konservativitätseigenschaften und damit die Qualität der numerischen Ergebnisse des TV–Filters sehr stark mit der Wahl der Filterparameter  $\varepsilon$  und  $\lambda$  variieren, ist der Filter für die Anwendung nach jedem Zeitschritt eines Differenzenverfahrens höherer Ordnung ungeeignet.

Seine Anwendung auf stationäre Lösungen von Differenzenverfahren höherer Ordnung führt dagegen selbst für die zweidimensionalen Eulergleichungen zu guten Ergebnissen, die für keine Parameterkombination Verschiebungsfehler aufweisen.

Um die Ursache für die unterschiedlichen Ergebnisse der Filteranwendungen zu verstehen, betrachten wir wieder die Lax–Wendroff–Lösung für das Cauchy–Problem (1.49) auf einem groben Gitter aus 40 Gitterpunkten. Abbildung 3.16 vergleicht die ungefilterte stationäre Lax–Wendroff–Lösung zur Zeit  $t = 2.5$  bei einer CFL–Zahl von 0.9 mit der gefilterten Lösung, die sich bei Anwendung von 80 Filteriterationen auf die stationäre Lax–Wendroff–Lösung ergibt. Die Filterparameter sind  $\lambda = 1$  und  $\varepsilon = 10^{-8}$ , und es ist, auch bei jeglicher anderer Wahl der Parameter, kein Verschiebungsfehler zu beobachten. Allerdings stellen wir fest, daß die

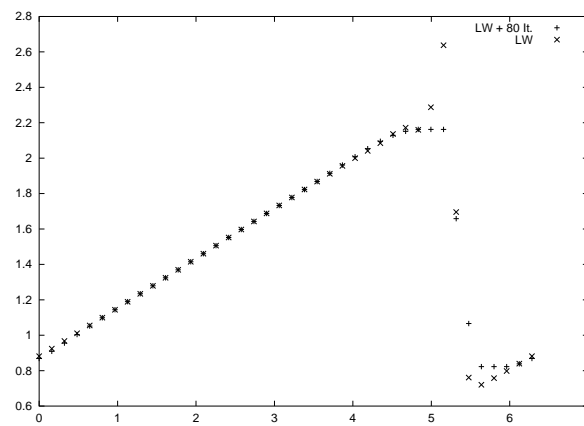


Abbildung 3.16: Anwendung von 80 Filteriterationen auf die stationäre Lax–Wendroff–Lösung für  $\lambda = 1$  und  $\varepsilon = 10^{-8}$  im Vergleich zur ungefilterten Lösung

Anwendung der Filterroutine dazu führt, daß eine größere Anzahl von Datenpunkten im Stoß liegen und darüber hinaus die Stoßhöhe verkleinert ist. Wird kein weiterer Lax–Wendroff–Schritt angewendet, ändert dies natürlich nichts an der Stoßlage. Führen wir dagegen für diesen Datensatz einen weiteren Lax–Wendroff–Schritt aus, ändert sich das Ergebnis des Differenzenverfahrens gegenüber der ungefilterten Lösung und die Konvergenz gegen die exakte Lösung von (1.49) ist nicht mehr gewährleistet.

Insgesamt zeigen die Anwendungen des TV–Filters von Chan, Osher und Shen auf numerische Lösungen hyperbolischer Erhaltungsgleichungen, daß es prinzipiell möglich ist, im Kontext der Bildverarbeitung entwickelte Filterroutinen als nichtlineare Diffusionsterme zu interpretieren und sie in Differenzenverfahren höherer Ordnung zu integrieren. Der Filter ist in der Lage, die Oszillationen stationärer Lösungen von Erhaltungsgleichungen zu glätten, ohne dabei Höhe oder Lage von Unstetigkeiten zu zerstören. Mit minimalem zusätzlichem Rechenaufwand haben wir beispielsweise die Lax–Wendroff–Lösung für die zweidimensionalen Eulergleichungen verbessert.

Bei der Filteranwendung nach jedem Zeitschritt des Zentrale–Differenzen–Verfahrens zweiter Ordnung konnten wir die Fähigkeit des Filters, Oszillationen zu glätten und instabile Verfahren zu stabilisieren, ebenfalls demonstrieren, aber es hat sich auch gezeigt, daß der Filter, abhängig von der Wahl der Filterparameter, Verschiebungsfehler in der numerischen Lösung erzeugen kann. Diese Fehler lassen sich darauf zurückführen, daß der Filter nur für gewisse Parameterkombinationen konservativ ist, und sich damit nicht für die Anwendung nach je-

dem Zeitschritt eines Differenzenverfahrens eignet.

Die Untersuchung eines weiteren TV-Filters der Bildverarbeitung im nächsten Unterkapitel soll Aufschluß darüber geben, ob andere Diskretisierungen der gleichen Modellgleichung zu ähnlichen Problemen und Eigenschaften führen, wie wir sie für den Filter von Chan, Osher und Shen festgestellt haben.

## 3.2 TV-Filter von Marquina und Osher

Im Folgenden untersuchen wir einen weiteren TV-Filter der Bildverarbeitung, der von Marquina und Osher entwickelt wurde und auf dem gleichen Minimierungsproblem und damit auf dem gleichen kontinuierlichen Modell beruht wie der Filter von Chan, Osher und Shen.

### 3.2.1 Grundlegende Gleichungen und kontinuierliches Filtermodell

Marquina und Osher betrachten bei der Herleitung ihres Filtermodells in [42] ein verrauschtes und gleichzeitig verschwommenes Bild  $\tilde{u}$ , aus dem sie das tatsächliche Bild  $u$  wiederherstellen wollen. Beide Bilder hängen dabei gemäß des Modells

$$\tilde{u} = k * u + n$$

zusammen, wobei  $k * u$  die Faltung von  $u$  und  $k$  bezeichnet und  $k$  eine Kernfunktion ist. Der additive Term  $n$  stellt Gaußsches weißes Rauschen dar, so daß die Werte von  $n$  an den verschiedenen Bildpunkten  $i$  unabhängige Zufallsvariablen sind, die jeweils eine Gaußverteilung mit Erwartungswert Null und Varianz  $\sigma^2$  besitzen. Die Faltung von  $u$  mit einer Kernfunktion  $k$  stellt die Unschärfe des Bildes  $\tilde{u}$  dar, die beim Filterprozeß das Aufsteilen von Kanten notwendig macht. Da in unseren oszillierenden numerischen Lösungen von Erhaltungsgleichungen keine Kantenverschärfung erreicht werden muß, können wir für unsere Anwendungen die Faltung mit einer Kernfunktion außer Betracht lassen und das zugrundeliegende Modell von Anfang an auf

$$\tilde{u} = u + n$$

reduzieren. Zusammen mit den Bedingungen an  $n$  stimmt diese Modellvorstellung genau mit der von Chan, Osher und Shen überein. Ziel ist es nun, das tatsächliche Bild  $u$ , ausgehend von seiner verrauschten Version  $\tilde{u}$ , zu berechnen, indem Informationen über das Rauschen und weiteres Vorwissen über das Bild, wie Glattheit oder Existenz von Kanten, ausgenutzt werden. Die Vorkenntnisse über das reale Bild werden durch ein Funktional  $R(u)$  in das zu lösende Variationsproblem einbezogen.  $R(u)$  ist dabei ein Maß für die Bildqualität von  $u$ , in dem Sinne, daß kleinere Werte von  $R(u)$  besseren Bildern entsprechen. Der Prozeß zur Berechnung von  $u$  besteht aus der Auswahl des qualitativ besten Bildes aus denen, die die Bedingungen an das statistische weiße Rauschen erfüllen, wobei die Qualität des Bildes durch die Bewertungsfunktion  $R(u)$  gemessen wird. Beispiele für Funktionale  $R(u)$ , die sich in der Literatur finden lassen ([67], [22]), sind  $R(u) = \|\Delta u\|_{L^2}$  und  $R(u) = \|\nabla u\|_{L^2}$ , wobei  $\Delta$  den Laplace-Operator und  $\nabla$  den Gradienten bezeichnet. Da die  $L^2$ -Norm des Gradienten zwar

Oszillationen entfernt, aber gleichzeitig die Gradienten, die den Kanten des Bildes entsprechen, zu stark glättet, verwenden Marquina und Osher als Funktional die TV-Norm

$$TV(u) = \int_{\Omega} |\nabla u| dx = \int_{\Omega} \sqrt{u_x^2 + u_y^2} dx,$$

die die Gradienten der Bildkanten erhält. Dabei bezeichnet  $\Omega$  das zugrundegelegte Rechengebiet, das eine beschränkte und offene Teilmenge des  $\mathbb{R}^2$  darstellt. Zur Vereinfachung sei  $\Omega$  im Folgenden stets das Einheitsquadrat.

Der gängige Lösungsansatz für die Berechnung von  $u$  ist dann durch das bedingte Variationsproblem

$$\min_u \int_{\Omega} |\nabla u| dx \quad \text{unter} \quad \frac{1}{2} \left( \int_{\Omega} (u - \tilde{u})^2 dx - |\Omega| \sigma^2 \right) = 0$$

gegeben. Dies entspricht genau dem Minimierungsproblem von Chan, Osher und Shen, so daß wir die zugehörige Euler-Lagrange-Gleichung sofort hinschreiben können,

$$0 = -\nabla \cdot \left( \frac{\nabla u}{|\nabla u|} \right) + \lambda(u - \tilde{u}). \quad (3.14)$$

Da diese Gleichung für Punkte mit  $\nabla u = 0$  nicht wohldefiniert ist, wird eine Regularisierung eingeführt, indem die Totalvariation durch

$$\int_{\Omega} |\nabla u|_{\beta} dx = \int_{\Omega} \sqrt{|\nabla u|^2 + \beta} dx$$

mit kleinem Parameter  $\beta > 0$  und der Bezeichnung

$$|x|_{\beta} = \sqrt{x^2 + \beta}, \quad |v|_{\beta} = \sqrt{|v|^2 + \beta} \quad \text{für } x \in \mathbb{R}, v \in \mathbb{R}^2$$

modifiziert wird.

Das zu Gleichung (3.14) zugehörige zeitabhängige Modell läßt sich analog durch

$$u_t = -\lambda(u - \tilde{u}) + \nabla \cdot \left( \frac{\nabla u}{|\nabla u|} \right) \quad (3.15)$$

darstellen. Dabei bezeichnet  $\tilde{u} = u(x, y, 0)$  das verrauschte Anfangssignal, und es gelten homogene Neumann-Randbedingungen, so daß  $\frac{\partial u}{\partial n} = 0$  am Rand des Definitionsbereiches erfüllt wird. Für wachsendes  $t$  nähern wir uns einer wiederhergestellten Version des tatsächlichen Bildes, wobei der zeitliche Prozeß Kantenerkennung und Glättung von Oszillationen gewährleisten soll. Beim Lösungsvorgang werden parabolische Gleichungen gelöst, daß heißt, Zeitschrittweite  $\Delta t$  und Raumschrittweite  $\Delta x$  hängen über die parabolische CFL-Bedingung

$$\frac{\Delta t}{\Delta x^2} \leq c |\Delta u|$$

mit festem  $c > 0$  miteinander zusammen. Dieser parabolische Zusammenhang bewirkt, daß eine stationäre Lösung nur mit sehr kleiner Zeitschrittweite erreicht werden kann und das resultierende Verfahren sehr viel Rechenzeit kostet.

Um dieses Problem zu umgehen, führen Marquina und Osher, ausgehend von (3.15), ein neues Modell ein. Sie multiplizieren Gleichung (3.14) mit dem Betrag des Gradienten von  $u$  und erhalten

$$u_t = -|\nabla u| \lambda(u - \tilde{u}) + |\nabla u| \nabla \cdot \left( \frac{\nabla u}{|\nabla u|} \right) \quad (3.16)$$

als neues Zeitschrittverfahren. Wie vorher werden homogene Neumann-Randbedingungen verwendet sowie ein a priori Schätzer für den Lagrangeschen Multiplikator  $\lambda$ . Analytisch betrachtet nähert sich dieses Verfahren dem gleichen stationären Zustand wie (3.15), so lange  $|\nabla u| \neq 0$  gilt, erreicht ihn aber weitaus schneller, da die Multiplikation mit  $|\nabla u|$  den parabolischen Term  $\nabla \cdot \left( \frac{\nabla u}{|\nabla u|} \right)$  nichtlinear regularisiert, denn die Glättungseffekte werden an den Stellen, an denen große Gradienten auftreten, durch den Multiplikator  $|\nabla u|$  verstärkt. In der Bildverarbeitung ist diese regularisierte Diffusion wohlbekannt. Der Diffusionsprozeß

$$u_t = |\nabla u| \nabla \cdot \left( \frac{\nabla u}{|\nabla u|} \right)$$

ist ein oft verwendetes Verfahren, das *mean curvature motion* (MCM) heißt. Für unsere Verwendungszwecke ist es wichtig, darauf hinzuweisen, daß diese Gleichung aufgrund des zusätzlichen Faktors  $|\nabla u|$  nicht mehr konservativ ist, was daher ebenso für das Verfahren (3.16) von Marquina und Osher gilt. Obwohl die Konservativität bei Anwendungen auf dem Gebiet der Erhaltungsgleichungen eine große Rolle spielt, wollen wir im Folgenden die Diskretisierung und numerische Anwendung von (3.16) untersuchen, um festzustellen, ob auch hier Verschiebungsprobleme in den Lösungen auftreten.

### 3.2.2 Diskretisierung und numerische Umsetzung

#### Diskretisierung der Modellgleichung

Der Einfachheit halber betrachten wir die Diskretisierung und Numerik von (3.16) im Eindimensionalen.

In 1D ist die Euler-Lagrange-Gleichung durch

$$0 = - \left( \frac{u_x}{|u_x|} \right)_x + \lambda(u - \tilde{u})$$

und zugehöriger regularisierter Version

$$0 = - \left( \frac{u_x}{|u_x|^\beta} \right)_x + \lambda(u - \tilde{u})$$

mit  $\beta > 0$  als Regularisierungsparameter und der Bezeichnung  $|u_x|_\beta = \sqrt{|u_x|^2 + \beta}$  gegeben. Darüber hinaus ist es möglich, die Euler-Lagrange-Gleichung mit Hilfe der Diracschen  $\delta$ -Funktion gemäß

$$0 = -\delta(u_x)u_{xx} + \lambda(u - \tilde{u})$$

darzustellen. Damit besitzt das entsprechende Zeitschrittmodell in 1D die Form

$$u_t = -|u_x|\lambda(u - \tilde{u}) + |u_x|\delta(u_x)u_{xx},$$

und mit

$$\begin{aligned}
|u_x| \left( \frac{u_x}{|u_x|^\beta} \right)_x &= |u_x| \left( \frac{u_x}{\sqrt{u_x^2 + \beta}} \right)_x \\
&= |u_x| \left( \frac{u_{xx}}{\sqrt{u_x^2 + \beta}} - \frac{u_x^2 u_{xx}}{(\sqrt{u_x^2 + \beta})^3} \right) \\
&= u_{xx} \left( \frac{|u_x|}{|u_x|^\beta} - \frac{u_x^2 |u_x|}{(u_x^2 + \beta) |u_x|^\beta} \right) \\
&\approx u_{xx} \left( 1 - \frac{u_x^2}{\beta + u_x^2} \right) = u_{xx} \frac{\beta}{\beta + u_x^2}
\end{aligned}$$

läßt es sich zu

$$u_t = -|u_x| \lambda(u - \tilde{u}) + \frac{\beta}{\beta + u_x^2} u_{xx}.$$

umschreiben. Für praktische Anwendungen verwenden Marquina und Osher die Approximation

$$\delta(z) \approx \beta \cdot (z^2 + \beta)^{-\frac{3}{2}}$$

als Näherung an die  $\delta$ -Funktion.

Da wir bei der numerischen Anwendung nicht mehr davon ausgehen können, daß das gegebene Signal überall stetig ist, stellt sich die Frage, in wieweit obige Umformungen auch für unstetiges  $u$  beziehungsweise unstetiges  $u_x$  gelten. Zu prüfen ist daher die Gültigkeit von

$$\left( \frac{u_x}{|u_x|} \right)_x = (\operatorname{sgn}(u_x))_x \stackrel{?}{=} \delta(u_x) u_{xx},$$

da die gewöhnliche Kettenregel nicht für unstetiges  $u_x$  gilt.

Für eine regularisierte Version der Signumfunktion erhalten wir

### Theorem 3.2

Sei

$$\widetilde{\operatorname{sgn}}_m(x) := \frac{2}{\pi} \arctan(mx) \quad (3.17)$$

eine regularisierte Version der Signumfunktion mit Regularisierungsparameter  $m \in \mathbb{N}$ , und sei  $u_x(x) \in BV$ . Dann gilt

$$\frac{\partial}{\partial x} \widetilde{\operatorname{sgn}}_m(u_x) = \delta_m(u_x) u_{xx}. \quad (3.18)$$

**Bemerkung:** Der Raum  $BV$  ist der Raum der auf  $\mathbb{R}^n$  definierten und summierbaren Funktionen, deren erste verallgemeinerte Ableitungen Maße sind.

Für den Beweis dieses Theorems verwenden wir die von Vol’pert und Hudjaev [71] entwickelte Kettenregel für unstetige Funktionen, die wir im Folgenden zusammen mit den notwendigen Definitionen darstellen. Nähere Erläuterungen zu verallgemeinerten Ableitungen und dem Raum  $BV$  sind ebenfalls in [71] oder in [72] zu finden.



**Definition 3.3**

Ein Punkt  $x_0$  heißt regulärer Punkt einer in einer Umgebung von  $x_0$  definierten Funktion  $u(x)$ , wenn es einen Einheitsvektor  $a \in \mathbb{R}^n$  gibt, so daß die approximativen Grenzwerte

$$u_a(x_0) := \lim_{x \rightarrow x_0, x \in \Pi_a(x_0)} u(x) \quad \text{und} \quad u_{-a}(x_0)$$

existieren. Dabei bezeichnet  $\Pi_a(x_0)$  die Halbebene  $(x - x_0, a) > 0$ , und  $a$  heißt definierender Vektor.

Für eine auf einer Menge  $G \subseteq \mathbb{R}^p$  definierte vektorwertige Funktion  $u(x) = (u_1(x), \dots, u_p(x))$  und eine auf  $\mathbb{R}^p$  definierte Funktion  $f$ , die für alle  $u$  stetig sei, erklären Vol'pert und Hudjaev die *gemittelte Verknüpfung* (averaged superposition) von  $f$  und  $u$ :

**Definition 3.4**

Sei  $x$  ein regulärer Punkt der vektorwertigen Funktion  $u$  und  $a$  der definierende Vektor. Dann heißt

$$\bar{f}(u(x)) = \int_0^1 f(u_a(x)t + u_{-a}(x)(1-t)) dt \quad (3.19)$$

gemittelte Verknüpfung der Funktionen  $f$  und  $u$ .

Mit diesen Definitionen läßt sich die Kettenregel für unstetiges  $u$  formulieren:

**Theorem 3.5**

Sei  $f(u)$  eine für alle  $u \in \mathbb{R}^p$  definierte, stetige und stetig differenzierbare Funktion. Sei  $u(x) = (u_1(x), \dots, u_p(x)) \in BV$ , also eine Funktion beschränkter Variation, und sei  $\bar{f}_{u_k}(u(x)) := \partial \bar{f}(u) / \partial u_k$  lokal summierbar bezüglich des Maßes  $\partial u_k / \partial x_i$  ( $k = 1, \dots, p$ ,  $i = 1, \dots, n$ ). Dann ist  $f(u(x)) \in BV$  und es gilt

$$\frac{\partial f(u(x))}{\partial x_i} = \sum_{k=1}^p \bar{f}_{u_k}(u(x)) \frac{\partial u_k}{\partial x_i} \quad \text{für } i = 1, \dots, n. \quad (3.20)$$

**Bemerkung**

- (a) Ist  $u(x)$  eine beschränkte vektorwertige Funktion, so gilt das Theorem auch, wenn  $\bar{f}_{u_k}(u(x))$  nicht lokal summierbar ist.
- (b) Ersetzen wir in (3.20)  $\bar{f}_{u_k}(u(x))$  durch  $f_{u_k}(u(x))$ , so ergibt sich die bekannte Kettenregel für stetig differenzierbare Funktionen  $f$  und  $u$ .

Mit Hilfe von Theorem 3.5 können wir jetzt das Theorem 3.2 beweisen:

**Beweis** Aus (3.17) ergibt sich zunächst für die Dirac-Funktion die Regularisierung

$$\delta_m(x) = \frac{2}{\pi} \frac{m}{1 + m^2 x^2}.$$

Darüber hinaus liefert uns (3.20) für  $p = i = 1$

$$\frac{\partial f(u(x))}{\partial x} = \bar{f}_u(u(x)) \frac{\partial u}{\partial x},$$

so daß sich für  $f(u_x(x)) = \widetilde{\text{sgn}}_m(u_x(x))$  die Ableitung

$$\frac{\partial}{\partial x} \widetilde{\text{sgn}}_m(u_x) = (\overline{\widetilde{\text{sgn}}_m})_{u_x}(u_x(x)) \frac{\partial u_x}{\partial x} \quad (3.21)$$

ergibt, wobei  $\partial u_x / \partial x = u_{xx}$  als Maß zu interpretieren ist.

Mit (3.19) und

$$\frac{\partial}{\partial u_x} (\widetilde{\text{sgn}}_m(u_x(x))) = \frac{\partial}{\partial u_x} \left( \frac{2}{\pi} \arctan(mu_x) \right) = \frac{2}{\pi} \frac{m}{1 + m^2 u_x^2}$$

erhalten wir

$$\begin{aligned} (\overline{\widetilde{\text{sgn}}_m})_{u_x}(u_x(x)) &= \int_0^1 \frac{2}{\pi} \frac{m}{1 + m^2(u_{x,a}t + u_{x,-a}(1-t))^2} dt \\ &= \int_0^1 \frac{2}{\pi} \frac{m}{1 + m^2((u_{x,a} - u_{x,-a})t + u_{x,-a})^2} dt \\ &= \frac{2m}{\pi} \int_0^1 \frac{1}{m^2(u_{x,a} - u_{x,-a})^2 t^2 + 2m^2(u_{x,a} - u_{x,-a})u_{x,-a}t + 1 + m^2 u_{x,-a}^2} dt \\ &= \frac{2m}{\pi} \left[ \frac{2}{\sqrt{4m^2(u_{x,a} - u_{x,-a})^2}} \arctan((u_{x,a} - u_{x,-a})t + mu_{x,-a}) \right]_0^1 \\ &= \frac{2}{\pi} \frac{\arctan(mu_{x,a}) - \arctan(mu_{x,-a})}{u_{x,a} - u_{x,-a}} \\ &= \frac{2}{\pi} (\arctan(mu_x))' = \frac{2}{\pi} \frac{m}{1 + m^2 u_x^2} = \delta_m(u_x), \end{aligned}$$

und zusammen mit (3.21) ergibt sich die Behauptung

$$\frac{\partial}{\partial x} \widetilde{\text{sgn}}_m(u_x) = \delta_m(u_x) u_{xx}.$$

■

Damit ist gezeigt, daß die von Marquina und Osher vorgenommenen Umformungen auch für unstetiges  $u_x \in \text{BV}$  gelten, so daß ihr resultierendes Filtermodell für unstetige Signale anwendbar ist.

Im Folgenden beschreiben wir die von Marquina und Osher gewählte Diskretisierung der 1D-Modellgleichung

$$u_t = -|u_x| \lambda(u - \tilde{u}) + \frac{\beta}{\beta + u_x^2} u_{xx}. \quad (3.22)$$

Sei  $U_i^n$  die numerische Approximation von  $u(x_i, t_n)$  mit  $x_i = i\Delta x$  und  $t_n = n\Delta t$ , und sei  $\tilde{U}_i$  die Approximation an  $\tilde{u}(x_i)$ . Dann ist das Differenzenverfahren für (3.22) durch

$$\frac{U_i^{n+1} - U_i^n}{\Delta t} = -|u g_i| \lambda(U_i^n - \tilde{U}_i) + \frac{\beta}{\beta + g_i^2} \frac{U_{i+1}^n - 2U_i^n + U_{i-1}^n}{\Delta x^2} \quad (3.23)$$

gegeben, wobei

$$g_i = \frac{U_{i+1}^n - U_{i-1}^n}{2\Delta x}$$

gilt, und  $ug_i$  ist der *upwind*-Gradient ist, für den

$$ug_i = \frac{U_i^n - U_{i-1}^n}{\Delta x}$$

für  $g_i(U_i^n - \tilde{U}_i) > 0$  und

$$ug_i = \frac{U_{i+1}^n - U_i^n}{\Delta x}$$

für  $g_i(U_i^n - \tilde{U}_i) < 0$  gilt.

## Numerische Testfälle

### Filteranwendung auf stationäre Lösungen

Wir beginnen die numerischen Testfälle, indem wir ein erstes Anwendungsbeispiel von Marquina und Osher nachahmen, um zu überprüfen, ob abhängig von den Filterparametern Verschiebungsfehler in der Lösung auftreten können. Da der Filter dabei auf eine stationäre Lösung angewendet wird, erwarten wir keine Verschiebungsfehler.

Wir betrachten ein 1D-Signal mit Werten im Bereich von  $[0, 255]$ , erzeugen ein weißes Rauschen und überlagern beides, so daß wir, ebenso wie Marquina und Osher, ein verrauschtes Signal mit *signal to noise ratio* (SNR)  $\approx 5$  erhalten, das zusammen mit dem unverrauschten Anfangssignal im linken Teil von Abbildung 3.17 dargestellt ist. Das verrauschte Signal dient uns als zu filternder Datensatz  $\tilde{u}$ , auf den wir 80 Iterationen des Differenzenschemas (3.23) anwenden. Die CFL-Zahl sei dabei  $\text{CFL} = 0.25$  und als Filterparameter wählen wir, wie Marquina und Osher,  $\beta = 15$  und  $\lambda = 0.005$ . Für  $\lambda$  wurde dabei der größtmögliche Wert gewählt, für den das Verfahren noch stabil ist, wenn die Zeitschrittweite durch ein explizites Euler-Schema berechnet wird. Der Parameter  $\beta$  ist abhängig von dem SNR-Wert, das heißt, von dem Verhältnis zwischen tatsächlichem Signal und Rauschen. Abbildung 3.17 rechts zeigt das entrauschte Signal, das wir nach Anwendung der Filteriterationen erhalten. Die Oszillationen sind weitgehend geglättet und die Lage des ursprünglichen Signals ist vollständig erhalten.

Wiederholen wir den beschriebenen Filtervorgang für andere Parameterwerte  $\lambda$  und  $\beta$ , erhalten wir zwar weniger entrauschte oder überglättete Lösungen, aber die Lage des jeweiligen Signals bleibt stets die ursprüngliche. Die iterative Anwendung des Differenzenverfahrens (3.23) auf ein vorgegebenes festes Eingangssignal liefert damit keinerlei Verschiebungsfehler in der Lösung, was für ein stationäres Signal zu erwarten war.

Vernachlässigen wir bei der Simulation den diskretisierten Regularisierungsfaktor  $|ug_i|$ , ergeben sich keine nennenswerten Änderungen in den gefilterten Lösungen.

Als nächstes testen wir die Funktionalität des diskreten Filters (3.23) bei Anwendungen im Kontext hyperbolischer Erhaltungsgleichungen.

Dazu betrachten wir zunächst die Anwendung der Filterroutine auf eine gegebene oszillierende Lösung. Grundlage dafür ist das einfache Cauchy-Problem (1.49) für Burgers' Gleichung, das wir auf einem Gitter aus 400 äquidistanten Gitterpunkten mit Hilfe des Lax-Wendroff-Verfahrens

$$\begin{aligned} U_i^{n+1} = U_i^n & - \frac{\lambda}{2}(f(U_{i+1}^n) - f(U_{i-1}^n)) \\ & + \frac{\lambda^2}{2} \left[ \frac{1}{2}(U_i^n + U_{i+1}^n)(f(U_{i+1}^n) - f(U_i^n)) - \frac{1}{2}(U_{i-1}^n + U_i^n)(f(U_i^n) - f(U_{i-1}^n)) \right] \end{aligned}$$

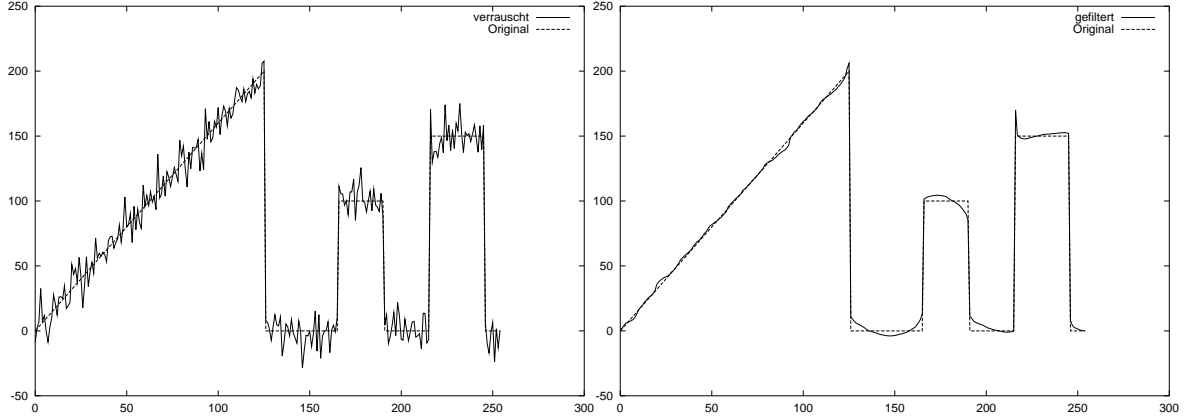


Abbildung 3.17: Originales und verrauschtes Signal (links) bzw. originales und wiederhergestelltes Signal (rechts)

mit  $f(u) = \frac{1}{2}u^2$  lösen. Auf die stationäre Lax–Wendroff–Lösung wenden wir mehrere Iterationen des diskreten Filters (3.23) an. Übernehmen wir dabei für die Filterschritte alle Filtereinstellungen vom nachprogrammierten Beispiel und auch die Größen der im Filter eingehenden Terme  $\Delta x$  und  $\Delta t$  (unabhängig von den entsprechenden Werten im Lax–Wendroff–Verfahren), also  $\Delta x = 1$ ,  $\Delta t = 0.3$ ,  $\lambda = 0.05$ ,  $\beta = 15$  und  $\text{CFL} = 0.25$ , so ergibt sich nach 10 Filteriterationen die in Abbildung 3.18 links dargestellte Lösung. Zum besseren Vergleich ist ebenfalls die reine Lax–Wendroff–Lösung ohne Filterroutine abgebildet. Die Oszillationen sind vollständig geglättet, wobei allerdings auch der tatsächliche Stoß etwas überglättet wurde. Die Lage des Stoßes ist exakt erhalten und ändert sich auch bei der Variation der Filterparameter nicht. Für  $\Delta x \rightarrow 0$  konvergiert die gefilterte Lösung.

Erhöhen wir die Anzahl der Iterationen, wird die gefilterte Lösung immer glatter, bis von dem Signal nur noch eine horizontale Linie übrig ist. Auch für größer werdendes  $\beta$  glättet der Filter immer stärker, und für  $\beta \rightarrow 0$  wird immer weniger gefiltert, bis die gefilterte Lösung irgendwann mit der ungefilterten übereinstimmt. Für kleiner werdendes  $\lambda$  ändert sich kaum etwas an der Lösung, für  $\lambda \geq 9.5$  entstehen allerdings zusätzliche Oszillationen und die Lösung wird instabil.

Da für die von Marquina und Osher übernommenen, festen Werte für  $\Delta x$  und  $\Delta t$  jegliche Motivierung fehlt, implementieren wir den Filteralgorithmus im Folgenden mit den tatsächlichen Werten für  $\Delta x$  und  $\Delta t$  aus dem Lax–Wendroff–Verfahren. Um die Konsistenz der gefilterten Lösung zu gewährleisten, müssen die Filterterme zusätzlich mit  $\Delta x$  multipliziert werden, so daß der Filter gemäß

$$\frac{U_i^{n+1} - U_i^n}{\Delta t} = -\Delta x |u g_i| \lambda (U_i^n - \tilde{U}_i) + \Delta x \frac{\beta}{\beta + g_i^2} \frac{U_{i+1}^n - 2U_i^n + U_{i-1}^n}{\Delta x^2} \quad (3.24)$$

modifiziert wird. Diese Filteranweisung implementieren wir, wie bei der vorangegangenen Anwendung, nach Beendigung des Lax–Wendroff–Verfahrens und erhalten eine gefilterte Lösung, die zusammen mit der zugehörigen Lax–Wendroff–Lösung im rechten Teil von Abbildung 3.18 dargestellt ist.

Die verwendeten Filterparameter sind  $\beta = 400$  und  $\lambda = 0.05$ , die Anzahl der Filteriterationen

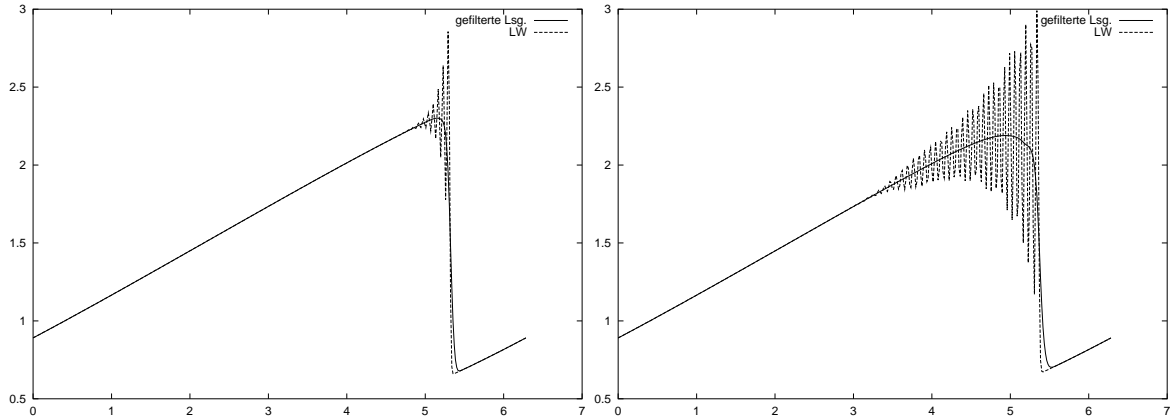


Abbildung 3.18: Mit (3.23) (links) bzw. (3.24) (rechts) gefilterte Lax-Wendroff-Lösung im Vergleich zur ungefilterten Lösung

beträgt 500, die CFL-Zahl ist 0.05 und gerechnet wurde auf 400 äquidistanten Gitterpunkten.

Da das Lax-Wendroff-Verfahren für kleine CFL-Zahlen stärker oszilliert, und die Filterleistung bei stärkeren Oszillationen besser dokumentiert werden kann, haben wir die CFL-Zahl an dieser Stelle klein gewählt, aber auch für größere CFL-Zahlen ergibt die Filteranwendung stabile und sogar weniger überglättete Ergebnisse.

Wie die rechte Darstellung in Abbildung 3.18 zeigt, sind die Oszillationen der Lax-Wendroff-Lösung vollständig geglättet, aber auch der tatsächliche Stoß ist überglättet. Die Lage des Stoßes wird exakt beibehalten und ändert sich auch bei der Variation der Filterparameter nicht. Bei steigender Anzahl der Filteriterationen wird die Lösung immer glatter und konvergiert, wie oben, gegen eine horizontale Gerade. Der Parameter  $\beta$  kann beliebig groß gemacht werden, wobei umso stärker gefiltert wird, je größer  $\beta$  wird, so daß für größeres  $\beta$  die Anzahl der Iterationen verringert werden kann. Ab einem Schwellenwert der Größenordnung 1500 für  $\beta$  lassen sich keine weiteren Änderungen in der Lösung feststellen. Für  $\beta \rightarrow 0$  wird immer weniger gefiltert und die gefilterte Lösung konvergiert gegen die ungefilterte. Bei der Variation von  $\lambda$  läßt sich beobachten, daß das Verfahren für zu große Werte von  $\lambda$  ( $\lambda > 10000$ ) instabil wird, da dann der erste Summand auf der rechten Seite der Filtergleichung (3.24) unverhältnismäßig groß wird und das Verfahren zerstört. Für  $\lambda \rightarrow 0$  ändert sich gegenüber der rechten Lösung aus Abbildung 3.18 nicht mehr viel, was verdeutlicht, daß die Komponente des Filters, die die Abweichung vom Ausgangsdatensatz bewertet, nur einen geringen Einfluß auf die Lösung hat. Für  $\Delta x \rightarrow 0$  konvergiert das resultierende Verfahren.

### Filteranwendung nach jedem Zeitschritt des Differenzenverfahrens

Im Gegensatz zu bisherigen Anwendungen wird der nichtlineare Filter jetzt nach jedem Zeitschritt des zugrundeliegenden Differenzenverfahrens implementiert.

Zunächst bleiben wir bei dem stabilen Lax-Wendroff-Verfahren und integrieren in dieses Schema nach jedem Zeitschritt einen Filterschritt. Aus Konsistenzgründen verwenden wir wieder den mit der Gitterweite  $\Delta x$  multiplizierten Filterausdruck (3.24). Als Filterparameter verwenden wir  $\beta = 15$ ,  $\lambda = 0.05$  und nach jedem Zeitschritt des Differenzenverfahrens fügen wir 3 Iterationen der Filterroutine (3.24) ein. Die CFL-Zahl beträgt 0.05 und  $\Delta x$  und

$\Delta t$  werden aus dem Zeitschrittverfahren übernommen. Das Ergebnis der Filterintegration in das Differenzenverfahren ist zusammen mit der ungefilterten Lax–Wendroff–Lösung im linken Teil von Abbildung 3.19 dargestellt. Die Oszillationen sind fast vollständig geglättet und die Lage des Stoßes ist erhalten. Für  $\Delta x \rightarrow 0$  konvergiert die Lösung gegen die tatsächliche und die Stoßlage bleibt weiterhin erhalten. Für größer werdendes  $\lambda$  verändert sich die Lösung zunächst kaum, erst bei sehr großen  $\lambda \approx 700000$  wird die Lösung instabil, und bei  $\lambda \rightarrow 0$  ändert sich ebenfalls kaum etwas an den numerischen Resultaten. Für größer werdendes  $\beta$  wird die Lösung glatter, für  $\beta \rightarrow 0$  wird immer weniger gefiltert, und in beiden Fällen bleibt die Stoßlage erhalten. Für  $0.1 \leq \beta < 12$  beobachten wir allerdings eine Verschiebung der Stoßlage nach links, die bei  $\beta = 4$  maximal wird und im rechten der Teil von Abbildung 3.19 dargestellt ist.

Auch für große CFL–Zahlen wie  $\text{CFL} = 0.98$  ist das resultierende Verfahren stabil, allerdings ist für die bisher verwendeten Parameter  $\beta = 15$  und  $\lambda = 0.05$  und einer CFL–Zahl größer oder gleich 0.1 eine geringe Verschiebung der Stoßlage nach rechts zu beobachten. Diese Verschiebung existiert aber nur für ein zu klein gewähltes  $\beta$ , ab  $\beta \geq 30$  stimmt die Stoßlage bei  $\text{CFL} = 0.98$  mit der physikalisch richtigen Lage überein. Reduzieren wir die Anzahl der Filterschritte von 3 auf einen pro Zeitschritt, ist die korrekte Stoßlage für die CFL–Zahl 0.98 ab  $\beta = 50$  gegeben. Für  $\beta = 200$  ergibt sich beispielsweise die in Abbildung 3.20 gegebene numerische Lösung, die für bessere Vergleichszwecke mit der zugehörigen Lax–Friedrichs–Lösung dargestellt ist. Größere Werte für  $\beta$  führen zu einer unnötig starken Glättung der Lösung, für kleinere Werte bleibt ein kleiner Überschwinger aus dem oszillierenden Lax–Wendroff–Verfahren erhalten.

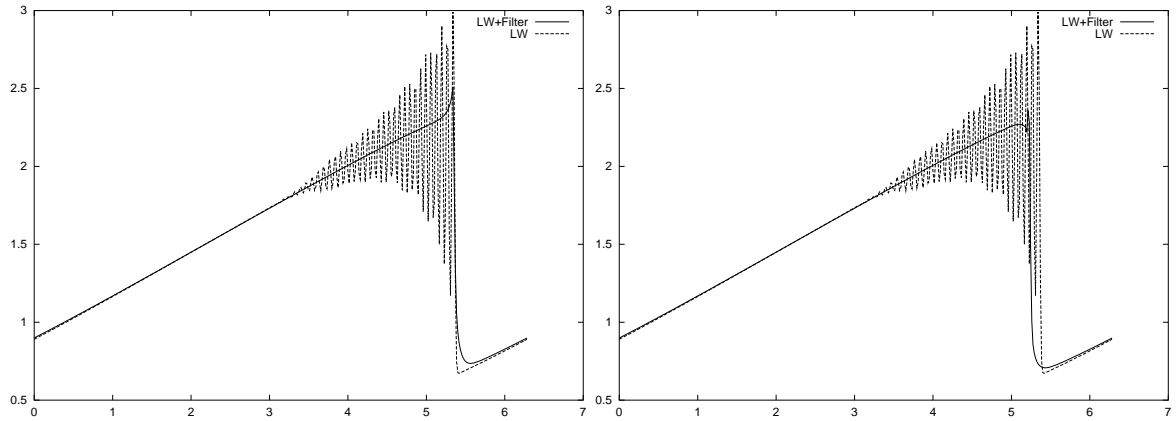


Abbildung 3.19: Vergleich zwischen ungefilterter und nach jedem Zeitschritt gefilterter Lax–Wendroff–Lösung, wobei das rechte Bild die maximale Stoßverschiebung bei  $\beta = 4$  darstellt

Da der in das Differenzenverfahren integrierte Filter fähig ist, rechentechnisch günstige, instabile Methoden zu stabilisieren, implementieren wir die Routine (3.24) als nächstes in das unbedingt instabile Zentrale–Differenzen–Verfahren

$$U_i^{n+1} = U_i^n - \frac{\Delta t}{2\Delta x}(U_{i+1}^n - U_{i-1}^n) \quad (3.25)$$

zweiter Ordnung. Abbildung 3.9 zeigt die oszillierende Zentrale–Differenzen–Lösung zum Zeitpunkt  $t = 2.5$  bei einer CFL–Zahl von 0.009 auf einem äquidistanten Gitter aus 400 Gitterpunkten.

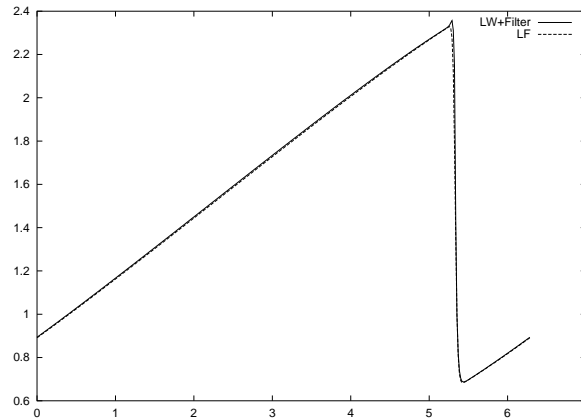


Abbildung 3.20: Mit einer Filteriteration pro Zeitschritt gefiltertes Lax–Wendroff–Verfahren bei CFL= 0.98

Fügen wir nach jeden Zeitschritt der zentralen Differenz drei Iterationen der Filterroutine (3.24) mit den Parametern  $\lambda = 0.05$ ,  $\beta = 15$  ein und verwenden als CFL–Zahl 0.05, so ergibt sich nach  $t = 2.5$  das in Abbildung 3.21 dargestellte Ergebnis. Die Oszillationen sind

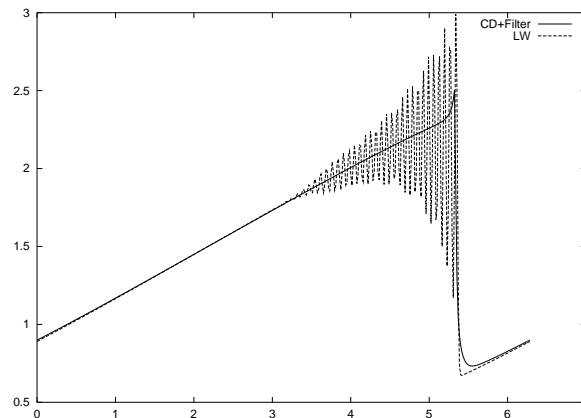


Abbildung 3.21: Mit drei Filteriterationen pro Zeitschritt gefiltertes Zentrales–Differenzen–Verfahren bei CFL= 0.05 im Vergleich zum Lax–Wendroff–Verfahren

fast vollständig geglättet, die Sprungschärfe ist erhalten und die Lage des Stoßes wird exakt wiedergegeben, was ein Vergleich mit der entsprechenden Lax–Wendroff–Lösung zeigt.

Für das Verhalten der Lösung bei Variation der Parameter  $\lambda$ ,  $\beta$  und der Anzahl der Filteriterationen gilt das gleiche wie im vorangegangenen Beispiel, mit der Ausnahme, daß jetzt das resultierende Verfahren für zu kleine Werte von  $\beta$  (ab 0.01) instabil wird, da das zugrundeliegende Differenzenverfahren nicht mehr stabil ist. Ebenso wie im letzten Beispiel tritt auch hier für  $0.1 \leq \beta < 12$  eine Verschiebung der Stoßlage nach links auf.

Das aus zentraler Differenz und Filterschritt zusammengesetzte Verfahren ist auch für größere CFL–Zahlen bis zu CFL= 0.5 stabil, wobei der Parameter  $\beta$  groß genug gewählt sein muß, damit keine Verschiebungsfehler in der Lösung auftreten. Bei drei Filteriterationen pro Zeitschritt und der CFL–Zahl 0.5 tritt ab  $\beta = 30$  kein Verschiebungsfehler mehr auf. Selbst bei nur einer Filteriteration pro Zeitschritt und CFL= 0.5 ergibt sich für  $\beta \geq 70$  ein stabiles Verfah-

ren, daß ab  $\beta = 500$  keinen Verschiebungsfehler mehr aufweist. Für  $\beta = 700$  ist die numerische Lösung, die sich bei CFL= 0.5 und einer Filteriteration nach jedem Zeitschritt der zentralen Differenz ergibt, in Abbildung 3.22 zusammen mit der zugehörigen Lax–Friedrichs–Lösung dargestellt. Dabei ist zu beachten, daß die Differenz  $U_i^n - \tilde{U}_i$  bei nur einer Filteriteration nach jedem Zeitschritt Null ist, so daß sich die Filtergleichung (3.24) zu

$$\frac{U_i^{n+1} - U_i^n}{\Delta t} = \Delta x \frac{\beta}{\beta + g_i^2} \frac{U_{i+1}^n - 2U_i^n + U_{i-1}^n}{\Delta x^2}$$

vereinfacht.

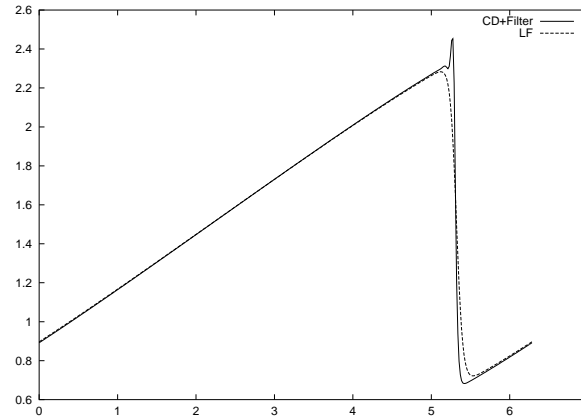


Abbildung 3.22: Mit einer Filteriteration pro Zeitschritt gefiltertes Zentrales–Differenzen–Verfahren bei CFL= 0.5 und  $\beta = 700$  im Vergleich zum Lax–Friedrichs–Verfahren

Das numerisch beobachtete Verhalten des aus zentraler Differenz und Filterschritt zusammengesetzten Verfahrens läßt sich auch anhand der Filtergleichung (3.24) erklären. Der Datensatz  $U^n$  in (3.24) bezeichnet das Ergebnis eines Zeitschritts mit dem Zentrale–Differenzen–Verfahren zweiter Ordnung (3.25), so daß der Filter von diesem Ergebnis am Gitterpunkt  $i$  den Term

$$\Delta t \Delta x |u g_i| \lambda (U_i^n - \tilde{U}_i)$$

abzieht und den Summanden

$$\frac{\Delta t}{\Delta x} \frac{\beta}{\beta + g_i^2} (U_{i+1}^n - 2U_i^n + U_{i-1}^n)$$

addiert. Der Term, der subtrahiert wird, ist in unseren Anwendungen sehr klein, was schon daran zu erkennen ist, daß sich die numerischen Ergebnisse für  $\lambda \rightarrow 0$  nur wenig ändern. Daher verändert hauptsächlich der zweite Filterterm das Ergebnis der zentralen Differenz, wobei der zusätzliche Summand proportional zur diskretisierten zweiten Ableitung ist und damit einen Dämpfungsterm darstellt. Für sehr großes  $\beta$  und  $\Delta t/\Delta x = 0.5$  ergibt sich beispielsweise

$$\frac{\Delta t}{\Delta x} \frac{\beta}{\beta + g_i^2} \approx \frac{1}{2},$$

so daß der Summand dem Diffusionsterm des Lax–Friedrichs–Verfahrens entspricht. Da für unsere Anwendungen allerdings

$$\frac{\Delta t}{\Delta x} \frac{\beta}{\beta + g_i^2} \ll \frac{1}{2}$$



gilt, ergibt das aus zentraler Differenz und Filterschritt zusammengesetzte Verfahren bessere Ergebnisse als das Lax–Friedrichs–Verfahren.

Um den Einfluß des Faktors  $|\nabla u|$ , den Marquina und Osher in ihr Filtermodell einfügen, numerisch zu untersuchen, wird im Folgenden die Anwendung des Filters nach jedem Zeitschritt der zentralen Differenz wiederholt, wobei der Filteralgorithmus jetzt ohne den Faktor  $|\nabla u|$  diskretisiert wird.

Aus dem Modell

$$u_t = -\lambda(u - \tilde{u}) + \delta(u_x)u_{xx}$$

erhalten wir analog zu (3.23) mit den obigen Bezeichnungen die Diskretisierung

$$\frac{U_i^{n+1} - U_i^n}{\Delta t} = -\lambda(U_i^n - \tilde{U}_i) + \frac{\beta}{(\beta + g_i^2)^{\frac{3}{2}}} \frac{U_{i+1}^n - 2U_i^n + U_{i-1}^n}{\Delta x^2}. \quad (3.26)$$

Diese Filteranweisung implementieren wir nach jedem Zeitschritt des Zentrale–Differenzen–Verfahrens bei einer CFL–Zahl von 0.05 und den Filterparametern  $\lambda = 0.05$  und  $\beta = 15$  auf einem Gitter aus 400 äquidistanten Gitterpunkten. Zu beobachten ist, daß nun mehrere Filteriteration pro Zeitschritt notwendig sind, um gleich gute numerische Resultate wie im letzten Beispiel zu erzielen. Abbildung 3.23 zeigt die gefilterte Zentrale–Differenzen–Lösung im Vergleich zur entsprechenden Lax–Wendroff–Lösung. Dieses Resultat stimmt qualitativ mit der in Abbildung 3.21 dargestellten Lösung überein, wobei jetzt allerdings 20 Filteriterationen pro Zeitschritt ausgeführt wurden. Die Multiplikation mit  $|\nabla u|$  bewirkt also, wie die Theorie besagt, eine schnellere Konvergenz der gefilterten Lösung. Fügt man nach jeden Zeitschritt des Differenzenverfahrens eine größere Anzahl Filteriterationen ein, so wird die gefilterte Lösung glatter. Bei einer zu kleinen Anzahl von Iterationen (kleiner 17), bildet sich ein kleiner Verschiebungsfehler der Stoßlage nach links aus. Für größeres  $\beta$  wird zunehmend weniger gefiltert, so daß für  $\beta > 10000$  viele Oszillationen nicht mehr geglättet werden. Für kleineres  $\beta$  ergeben sich anfangs kaum Änderungen, bei sehr kleinem  $\beta \approx 0.0001$  wird die Lösung instabil. Läßt man den Parameter  $\lambda$  kleiner werden, bleibt die Lösung annähernd gleich, was auch für größer werdendes  $\lambda$  gilt. Erst ab  $\lambda = 5000$  treten kleine Oszillationen in der Lösung auf, und bei  $\lambda = 500000$  wird die Lösung instabil. Die Stoßlage bleibt bei der Variation der beiden Filterparameter, bis auf minimale Abweichungen, konstant richtig.

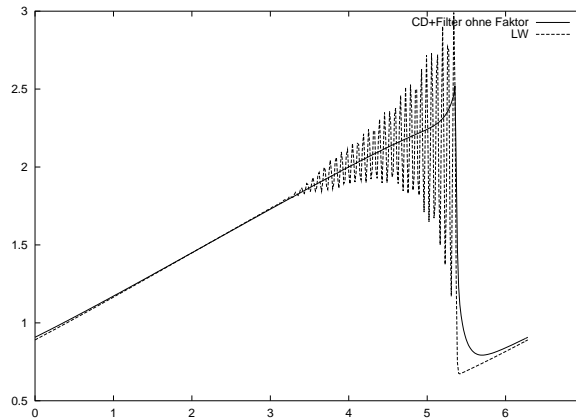


Abbildung 3.23: Mit (3.26) gefilterte zentrale Differenz

Die numerischen Resultate zeigen, daß die Anwendung des Marquina–Osher–Filters nach jedem Zeitschritt eines Differenzenverfahrens höherer Ordnung im Wesentlichen keine Verschiebung der tatsächlichen Stoßlage bewirkt, auch dann nicht, wenn der Term  $|\nabla u|$ , den Marquina und Osher künstlich in die Modellgleichung einfügen, entfernt wird. Die Qualität der gefilterten Lösungen hängt auch bei diesem Filter von der Wahl der Filterparameter ab. Allerdings wird die Lösung für hinreichend großes  $\beta$  bei der Variation der Parameter innerhalb eines sinnvollen Bereiches nicht zerstört oder durch Verschiebung der Stoßlage verfälscht, wie es bei dem TV–Filter von Chan, Osher und Shen der Fall ist. Die Variation der Parameter wirkt sich beim vorliegenden Filter hauptsächlich auf die Stärke seiner Glättungseigenschaften aus.

Obwohl die beiden untersuchten Filter auf dem gleichen kontinuierlichen Modell beruhen, weisen sie bei der Implementierung nach jedem Zeitschritt eines Differenzenverfahrens höherer Ordnung beträchtlich unterschiedliche Eigenschaften auf. Das unterschiedliche Verhalten der diskreten Filter zeigt, daß Eigenschaften der kontinuierlichen Modellgleichung bei der Diskretisierung verloren gehen und die diskreten Filter in ihrer Wirkungsweise stark von der gewählten Diskretisierung abhängen.

Im folgenden Unterkapitel entwickeln wir aus den kontinuierlichen Modellgleichungen, die den beiden Filtern zugrundeliegen, neue diskrete Filterroutinen und vergleichen ihre numerischen Eigenschaften bei der Implementierung in Differenzenverfahren höherer Ordnung.

### 3.3 Neue Diskretisierungen der Filtermodelle

Wie bereits bei der Herleitung der kontinuierlichen Filtergleichungen gezeigt, beruhen beide vorgestellten Filter auf dem gleichen bedingten Variationsproblem, für das die Euler–Lagrange–Gleichung durch

$$0 = \nabla \cdot \left( \frac{\nabla u}{|\nabla u|} \right) + \lambda(\tilde{u} - u) \quad (3.27)$$

gegeben ist, wobei

$$u_t = \nabla \cdot \left( \frac{\nabla u}{|\nabla u|} \right) + \lambda(\tilde{u} - u) \quad (3.28)$$

das zugehörige zeitabhängige Modell darstellt. Marquina und Osher ergänzen dieses Modell durch den Vorfaktor  $|\nabla u|$  und erhalten

$$u_t = |\nabla u| \nabla \cdot \left( \frac{\nabla u}{|\nabla u|} \right) + |\nabla u| \lambda(\tilde{u} - u). \quad (3.29)$$

Im Folgenden entwickeln wir für diese Gleichungen einfachere Modellvorstellungen, die wir direkt diskretisieren und auf die bereits bekannte Problemstellung anwenden.

### 3.3.1 Diskretisierungen für die Euler–Lagrange–Gleichung

Wir beginnen mit Gleichung (3.27). Sie läßt sich im Eindimensionalen zu

$$\partial_x \left( \frac{u_x}{|u_x|} \right) - \frac{u - \tilde{u}}{\frac{1}{\lambda}} = 0$$

umformen. Diese Darstellung legt es nahe,  $\frac{1}{\lambda}$  als  $\Delta t$  zu interpretieren, das heißt,  $\lambda$  wird als inverse Diffusionszeit aufgefaßt. Da weiterhin  $\frac{u_x}{|u_x|} = \text{sgn}(u_x)$  gilt, erhalten wir für die kontinuierliche Gleichung (3.27) die Darstellung

$$\partial_t u = \partial_x (\text{sgn}(u_x)). \quad (3.30)$$

Die direkte Diskretisierung dieses Modells liefert uns das numerische Verfahren

$$U_i^{n+1} = U_i^n + \frac{\Delta t}{\Delta x} (\text{sgn}(U_{i+1}^n - U_i^n) - \text{sgn}(U_i^n - U_{i-1}^n)), \quad (3.31)$$

das wir nach jedem Zeitschritt in das Zentrale–Differenzen–Verfahren zweiter Ordnung implementieren. Nach jedem Zeitschritt wenden wir fünf Filteriterationen an und berechnen die Lösung bei einer CFL–Zahl von 0.05 auf einem äquidistanten Gitter aus 400 Punkten. Die Anwendung von (3.31) nach jedem Zeitschritt der zentralen Differenz liefert zwar kein qualitativ gutes Ergebnis, da viele kleine Oszillationen selbst im linearen Anteil der Lösung auftreten, wie das linke Bild in Abbildung 3.24 im Vergleich mit der zugehörigen Lax–Wendroff–Lösung zeigt, aber unabhängig von der Anzahl der Filteriterationen pro Zeitschritt, ist nie ein Verschiebungsfehler der Stoßlage zu beobachten. Für  $\Delta x \rightarrow 0$  konvergiert das zusammengesetzte Verfahren, was die entsprechende Lösung, jetzt mit  $N = 3200$  Gitterpunkten berechnet, im rechten Teil der Abbildung 3.24 zeigt. Fast das gleiche Resultat läßt sich auch schon mit einer

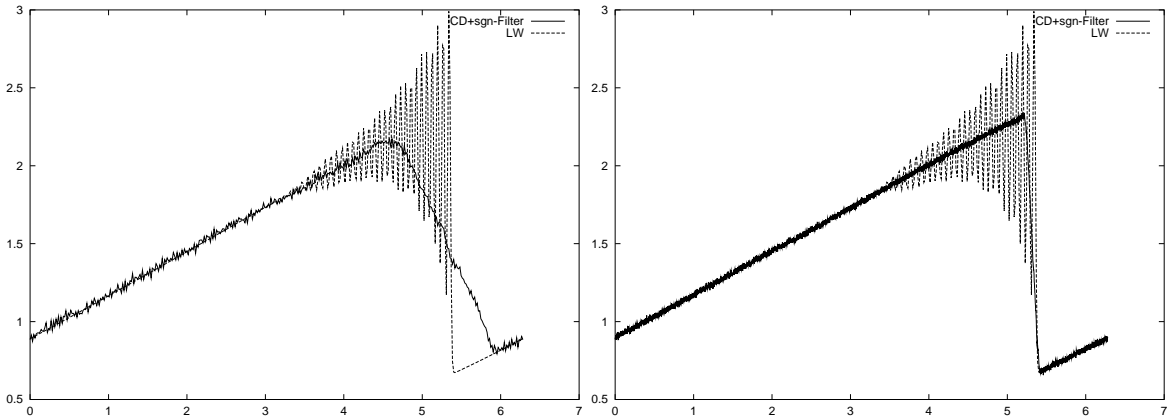


Abbildung 3.24: Anwendung des einfachen Signum–Filters (3.31) nach jedem Zeitschritt der zentralen Differenz für  $N = 400$  (links) und  $N = 3200$  (rechts)

Filteriteration pro Zeitschritt erreichen.

Die Konservativität der kontinuierlichen Filtergleichung ist aufgrund der Formulierung

$$\partial_t u = \partial_x \left( \frac{u_x}{|u_x|} \right),$$

der Divergenzform, zu erkennen, denn allgemein sind alle Gleichungen der Art

$$\partial_t u = f(u_x)_x = f'(u_x) \cdot u_{xx}$$

konservativ. Da auch die Diskretisierung konservativ ist, ist die Konservativität von (3.31) gewährleistet.

Die Qualität des Signum-Filters (3.31) läßt sich wesentlich verbessern, indem wir eine Gewichtung der Signum-Terme mit  $|U_{i+1} - U_i|$  beziehungsweise  $|U_i - U_{i-1}|$  einfügen, was ein gängiges Vorgehen bei der Filterregularisierung ist. Die diskrete Filtergleichung ist dann durch

$$U_i^{n+1} = U_i^n + \frac{\Delta t}{\Delta x} (|U_{i+1}^n - U_i^n| \operatorname{sgn}(U_{i+1}^n - U_i^n) - |U_i^n - U_{i-1}^n| \operatorname{sgn}(U_i^n - U_{i-1}^n)) \quad (3.32)$$

gegeben, und ihre Implementierung liefert für  $N = 400$  Gitterpunkte und 5 Iterationen pro Zeitschritt das im linken Teil von Abbildung 3.25 dargestellte Resultat. Es ist kein Verschiebungsfehler der Stoßlage zu beobachten, auch nicht bei mehreren oder weniger Iterationen pro Zeitschritt, das Verfahren konvergiert für  $\Delta x \rightarrow 0$  und liefert sehr glatte Ergebnisse. Durch die Anwendung von fünf Filterschritten pro Zeitschritt ist die Lösung etwas zu stark geglättet, was sich aber durch Reduktion der Filteriterationen verbessern läßt. Abbildung 3.25 zeigt rechts die entsprechende Lösung mit nur einer Filteriteration pro Zeitschritt. Das

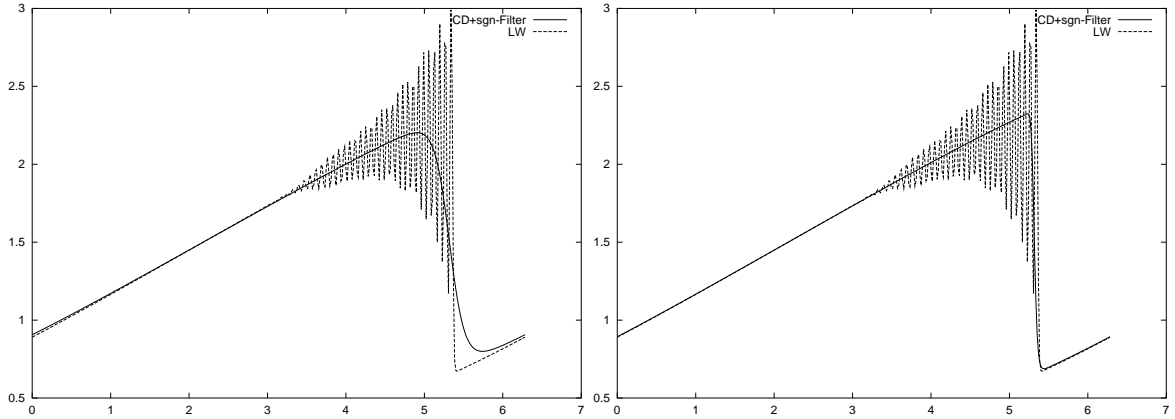


Abbildung 3.25: Anwendung von 5 Iterationen (links) bzw. einer Iteration (rechts) des regulierten Signum-Filters (3.32) nach jedem Zeitschritt der zentralen Differenz

zentrale Differenzenverfahren ist durch die Kombination mit der Filterroutine stabilisiert, die Oszillationen sind vollständig geglättet und Höhe und Lage des Stoßes sind fast vollständig erhalten.

Ausgehend von der kontinuierlichen Gleichung, die dem TV-Filter von Chan, Osher und Shen zugrundeliegt, haben wir damit einen neuen, sehr einfachen und schnellen nichtlinearen Filter konstruiert, der bei der Anwendung auf oszillierende Differenzenverfahren für Erhaltungsgleichungen sehr gute Ergebnisse erzeugt und dabei ganz ohne vorher einzustellende Parameter auskommt. Im Vergleich zu technisch aufwendigen und komplizierten Differenzenverfahren mit implizit definiertem Diffusionstensor haben wir einen rechentechnisch günstigen, diskreten Filter entwickelt, der Oszillationen glättet, ohne tatsächliche Stöße zu dämpfen, und als einfache, aber sehr geeignete Dissipation aufgefaßt werden kann.

Verschiebungsprobleme bei der Stoßlage treten bei dieser einfachen Diskretisierung, im Gegensatz zu der von Chan, Osher und Shen entwickelten Filterroutine, gar nicht erst auf.

Im Folgenden gehen wir noch kurz auf die Implementierung des einfachen Signum-Filters mit zusätzlichem Faktor  $\Delta x$  in den Filtertermen ein, wie es aus Konsistenzgründen beim Marquina-Osher-Filter gemacht wurde.

Multiplizieren wir die Filterterme von (3.31) mit  $\Delta x$ , ergibt sich

$$U_i^{n+1} = U_i^n + \Delta t \left( \text{sgn}(U_{i+1}^n - U_i^n) - \text{sgn}(U_i^n - U_{i-1}^n) \right), \quad (3.33)$$

und die Implementierung dieser Filterroutine führt zu einer Lösung mit Verschiebungsfehler der Stoßlage, wie das linke Bild von Abbildung 3.26 zeigt. Die Größe der Verschiebung ist

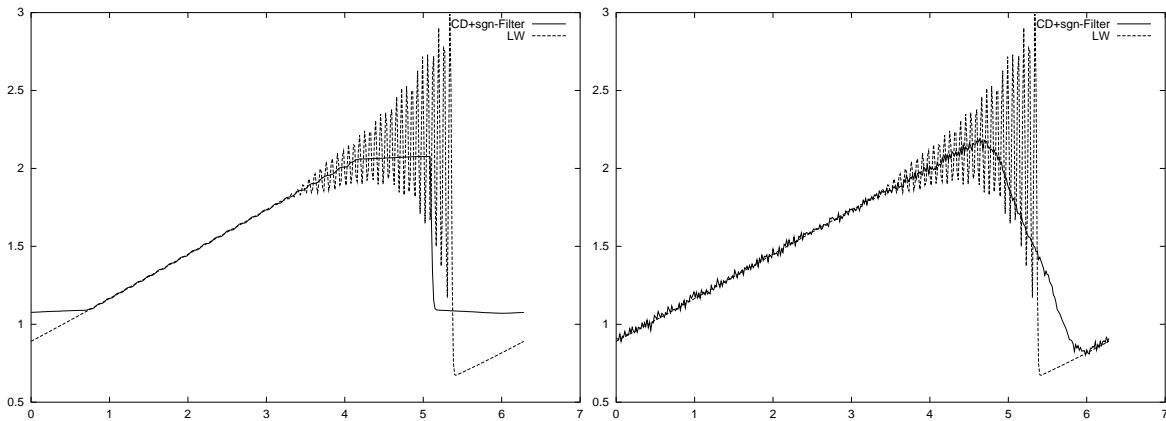


Abbildung 3.26: Anwendung des Signum-Filters (3.33) (links) und des zeitabhängigen Filters (3.34) (rechts) nach jedem Zeitschritt der zentralen Differenz

abhängig von der Anzahl der Filteriterationen pro Zeitschritt, bei sehr wenigen Iterationen ist noch kein Fehler zu erkennen, er entwickelt sich erst ab fünf Filterwiederholungen, wird dann mit zunehmender Iterationsanzahl größer und bleibt ab 10 Iterationen konstant. Für  $\Delta x \rightarrow 0$  nimmt der Verschiebungsfehler ab, was sich dadurch erklären läßt, daß die Filterterme jetzt mit  $\Delta x$  gewichtet sind, so daß die Filterwirkung und damit auch der Verschiebungsfehler mit  $\Delta x \rightarrow 0$  ebenfalls gegen Null geht.

Der Grund für die fälschlich erhöhten Werte am rechten und linken Rand der dargestellten Lösung ist die (globale) Konservativität der Filterroutine. Die Lösungswerte direkt vor dem Stoß sind zu weit nach unten korrigiert, so daß zum Ausgleich andere Werte zu weit nach oben korrigiert werden, damit die Konservativität eingehalten wird. Daß der Filter in der Tat konservativ ist, auch wenn sich lokal diskrete Situationen finden lassen, in denen der Filter lediglich einen Wert vermindert und keinen vergrößert, läßt sich, wie bei dem Verfahren (3.31), an seiner Divergenzform erkennen.

Das Experiment zeigt deutlich, daß allein das Einfügen der konstanten Größe  $\Delta x$  in den Filterterm einen Verschiebungsfehler in der Lösung hervorruft. Diese Tatsache ist bemerkenswert, da bei der Multiplikation der Filterterme mit Konstanten von der Anschauung her zwar eine quantitative Änderung der Lösung zu erwarten ist, ein qualitativ vollkommen anderes Verhalten aber nicht ohne Weiteres einsichtig ist.

### 3.3.2 Diskretisierungen für die zeitabhängige Modellgleichung

Für weitere Untersuchungen betrachten wir das zeitabhängige Modell (3.28) und wenden ebenso wie für Gleichung (3.27) eine einfache Diskretisierung mit Hilfe der Signumfunktion an.

Für den eindimensionalen Fall ergibt sich für (3.28) die direkte Diskretisierung

$$U_i^{n+1} = U_i^n + \frac{\Delta t}{\Delta x} (\operatorname{sgn}(U_{i+1}^n - U_i^n) - \operatorname{sgn}(U_i^n - U_{i-1}^n)) + \Delta t \cdot \lambda(\tilde{U}_i - U_i^n). \quad (3.34)$$

Die Anwendung dieser Filterroutine nach jedem Zeitschritt des Zentrale-Differenzen-Verfahrens liefert näherungsweise die gleichen Ergebnisse wie die diskretisierte Version (3.31) des zugehörigen stationären Modells, wie im rechten Teil von Abbildung 3.26 zu erkennen ist. Der Filterparameter für die dargestellte Lösung ist  $\lambda = 5$  und pro Zeitschritt wurden 10 Filteriterationen ausgeführt. Wieder sind viele kleine Oszillationen im linearen Lösungsanteil zu beobachten, aber ein Verschiebungsfehler der Stoßlage tritt nicht auf. Auch bei zunehmender Anzahl von Filteriterationen wird die Lösung zwar immer mehr geglättet, aber die richtige Stoßlage bleibt stets erhalten, was ebenso für die Variation des Parameters  $\lambda$  gilt, wobei der Filter für  $\lambda \rightarrow 0$  gegen die Filterroutine (3.31) konvergiert. Für wachsendes  $\lambda$  wird die Lösung flacher, bis sie für sehr großes  $\lambda > 3000$  instabil wird. Für  $\Delta x \rightarrow 0$  konvergiert das Verfahren ebenso wie (3.31).

Implementieren wir die Filterroutine (3.34) mit zusätzlichem Faktor  $\Delta x$  in den Filtertermen, erhalten wir fast identische Ergebnisse wie bei der Implementierung von (3.33), also glatte Lösungen mit Verschiebungsfehler. Die Ähnlichkeit zu den Ergebnissen der zugehörigen stationären Filtergleichung (3.33) ist offensichtlich, denn der Term  $\Delta t \lambda(\tilde{U}_i - U_i^n)$ , der sowieso wegen des Faktors  $\Delta t$  recht klein ist, wird zusätzlich mit  $\Delta x$  multipliziert, so daß sein Betrag noch kleiner und nicht mehr ausschlaggebend wird.

### 3.3.3 Diskretisierung der Modellgleichung von Marquina und Osher

Als dritte Variante diskretisieren wir das kontinuierliche Modell (3.29) von Marquina und Osher. Mit Hilfe der Signum-Formulierung ist die zugehörige diskrete Filterroutine im Eindimensionalen durch

$$\begin{aligned} U_i^{n+1} = & U_i^n + \frac{|U_{i+1}^n - U_i^n|}{\Delta x} \frac{\Delta t}{\Delta x} (\operatorname{sgn}(U_{i+1}^n - U_i^n) - \operatorname{sgn}(U_i^n - U_{i-1}^n)) \\ & + \frac{|U_{i+1}^n - U_i^n|}{\Delta x} \lambda(\tilde{U}_i - U_i^n) \end{aligned}$$

gegeben. Ihre Anwendung nach jedem Zeitschritt des Zentrale-Differenzen-Verfahrens liefert allerdings kein numerisch stabiles Verfahren, da der erste Filterterm

$$\frac{|U_{i+1}^n - U_i^n|}{\Delta x} \frac{\Delta t}{\Delta x} (\operatorname{sgn}(U_{i+1}^n - U_i^n) - \operatorname{sgn}(U_i^n - U_{i-1}^n))$$

aufgrund der Größe  $\Delta x^2$  im Nenner sehr groß werden kann und damit zur Instabilität führt. Abhilfe kann geschaffen werden, indem wir die Filterterme mit  $\Delta x$  multiplizieren, so daß der

Filter die Form

$$U_i^{n+1} = U_i^n + |U_{i+1}^n - U_i^n| \frac{\Delta t}{\Delta x} (\operatorname{sgn}(U_{i+1}^n - U_i^n) - \operatorname{sgn}(U_i^n - U_{i-1}^n)) + |U_{i+1}^n - U_i^n| \lambda (\tilde{U}_i - U_i^n) \quad (3.35)$$

erhält. Bei seiner Implementierung in die zentrale Differenz ergibt sich ein stabiles Verfahren, die Oszillationen sind vollständig geglättet, aber die Lösung weist einen großen Verschiebungsfehler der Stoßlage auf, der weitgehend unabhängig vom Parameter  $\lambda$  und der Anzahl der Filteriterationen ist. Abbildung 3.27 zeigt die Lösung für  $\lambda = 10$  und fünf Filteriterationen

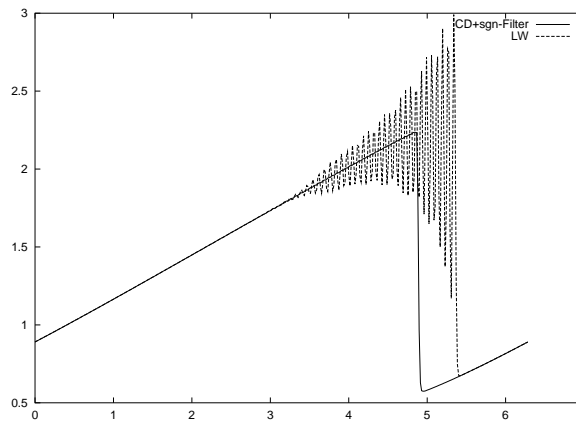


Abbildung 3.27: Anwendung des Filters (3.35) nach jedem Zeitschritt der zentralen Differenz

nen pro Zeitschritt.

Der Verschiebungsfehler kann dadurch begründet sein, daß der Ausdruck  $|u_x|$  nicht konservativ diskretisiert wurde, wobei zu beachten ist, daß schon die kontinuierliche Filtergleichung nicht mehr konservativ ist. Darüber hinaus kann auch die Multiplikation mit  $\Delta x$ , wie bei der Implementierung von (3.33), zu Verschiebungsfehlern führen.

### 3.3.4 Zeitabhängige Version des TV-Filters von Chan, Osher und Shen

Da der TV-Filter von Chan, Osher und Shen, der einem stationären Modell zugrundeliegt, Verschiebungsfehler produziert, hingegen der Filter von Marquina und Osher, der auf dem zugehörigen zeitabhängigen Modell beruht, im Wesentlichen keine Verschiebungsfehler hervorruft, läßt sich vermuten, daß der Verschiebungsfehler auch durch die Stationarität der Modellgleichung begründet sein kann. Daher formulieren wir die diskrete Filtergleichung von Chan, Osher und Shen im Folgenden so um, daß sie einem zeitabhängigen kontinuierlichen Modell entspricht.

Wir ergänzen die stationäre Gleichung

$$0 = -\lambda(u - \tilde{u}) + \nabla \cdot \left( \frac{\nabla u}{|\nabla u|} \right),$$

wie sie bei Chan und Mitarbeitern verwendet wird, durch die zeitliche Ableitung und erhalten das zeitabhängige Modell

$$u_t = -\lambda(u - \tilde{u}) + \nabla \cdot \left( \frac{\nabla u}{|\nabla u|} \right).$$

Den rechten Teil der Gleichung diskretisieren wir genauso wie Chan, Osher und Shen, und die linke Seite diskretisieren wir mit einer einfachen Vorwärtsdifferenz, so daß die neue Filterroutine durch

$$U_i^{n+1} = U_i^n - \Delta t U_i^n + \Delta t h_{ii-1} U_{i-1}^n + \Delta t h_{ii+1} U_{i+1}^n + \Delta t h_{ii} \tilde{U}_i \quad (3.36)$$

gegeben ist. Diesen Filteralgorithmus implementieren wir mit je 10 Iterationen und den Filterparametern  $\lambda = 0.001$  und  $\varepsilon = 10^{-8}$  nach jedem Zeitschritt des Zentrale-Differenzen-Verfahrens bei einer CFL-Zahl von 0.1. Das Ergebnis, verglichen mit der zugehörigen Lax-Wendroff-Lösung, ist in Abbildung 3.28 (links) dargestellt. Die Oszillationen im gefilterten Signal sind zwar noch nicht vollständig geglättet, aber der Verschiebungsfehler ist wesentlich geringer als vorher. Bei der Variation von  $\lambda$  bleibt die Stoßlage unverändert, für  $\lambda \rightarrow 0$  ändert sich die Lösung insgesamt nur noch wenig, für wachsendes  $\lambda$  nehmen die Oszillationen zu, aber die Stoßlage bleibt konstant. Ab  $\lambda > 10$  wird die Lösung instabil. Variieren wir dagegen die Anzahl der Filteriterationen, tritt allerdings bei mehreren Filteriterationen pro Zeitschritt eine größere Verschiebung der Stoßlage auf, die in Abbildung 3.28 rechts für 60 Iterationen dargestellt ist.

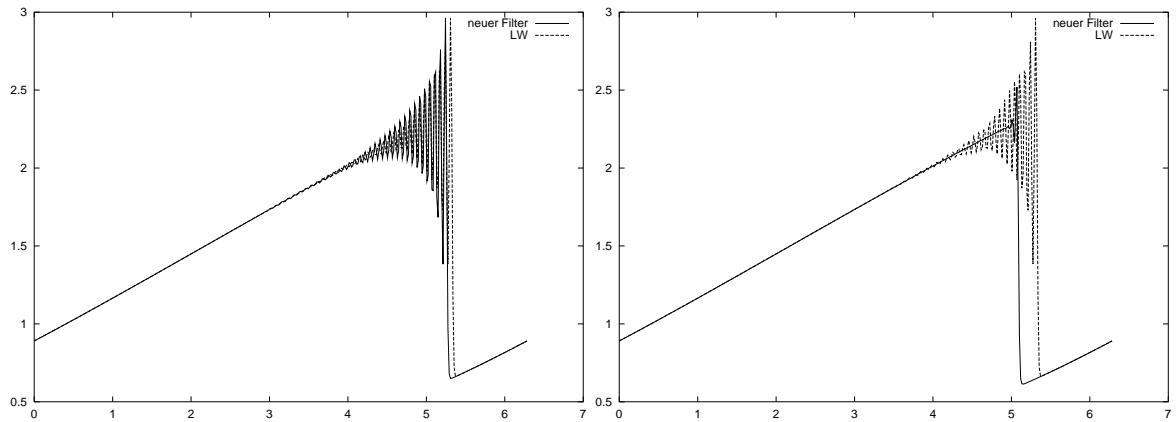


Abbildung 3.28: Anwendung der Filterroutine (3.36) nach jedem Zeitschritt des Zentrale-Differenzen-Verfahrens, implementiert mit 10 (links) bzw. 60 Filteriterationen

Damit haben wir gezeigt, daß wir den Verschiebungsfehler durch den Übergang vom stationären zum zeitabhängigen Modell bei Beibehaltung der Diskretisierung nicht beheben können, was wiederum die Vermutung verstärkt, daß der Verschiebungsfehler von der Diskretisierung und ihrer Nichtkonservativität abhängt.

### 3.3.5 Zusammenfassung der Beobachtungen

Abschließend betrachtet, zeigen die durchgeführten Experimente, daß sich die kontinuierlichen Modelle (3.27) und (3.28) mit Hilfe der Signumfunktion so umschreiben und diskretisieren



lassen, daß bei der numerischen Anwendung nach jedem Zeitschritt des zugrundeliegenden Differenzenverfahrens keine Verschiebungsfehler in der Lösung auftreten. Da die Diskretisierung, die von Chan, Osher und Shen gewählt wird, bei der gleichen Anwendung zu Fehlern bei der Stoßlage führt, bestätigt dies die Vermutung, daß die Verschiebungsfehler auf Diskretisierungsproblemen wie der Zerstörung der Konservativitätseigenschaft beruhen.

Darüber hinaus haben wir, aufbauend auf dem kontinuierlichen Modell (3.27), den parameterfreien Filter (3.32) entwickelt, der als einfache nichtlineare Dissipation das Zentrale-Differenzen-Verfahren zweiter Ordnung stabilisiert, die Oszillationen glättet und die Höhe und Lage der Unstetigkeit erhält.

Wie stark unterschiedliche Diskretisierungen des gleichen kontinuierlichen Modells die numerischen Eigenschaften der diskreten Filter beeinflussen, läßt sich auch an der Diskretisierung von Gleichung (3.29) erkennen. Implementieren wir den von Marquina und Osher entwickelten diskreten Filter mit dem zusätzlichen Faktor  $\Delta x$  nach jedem Zeitschritt des Differenzenverfahrens, ist kein Verschiebungsfehler in der numerischen Lösung zu beobachten. Wählen wir hingegen eine Diskretisierung mittels der Signumfunktion, tritt eine starke Stoßverschiebung auf, die auf die fehlende Konservativität des Filters zurückzuführen ist.

Mit Hilfe des letzten Experiments, der Implementierung von (3.36), können wir ausschließen, daß der Verschiebungsfehler von der Stationarität des zugrundegelegten Modells abhängt.

Abgesehen von der Feststellung, daß ungünstige Diskretisierungen Ursache für auftretene Stoßverschiebungen sind, zeigen die angeführten Experimente, daß allein die Multiplikation der Filterterme mit Konstanten für das Vorkommen von Verschiebungsfehlern verantwortlich sein kann, was der anschaulichen Vorstellung in erster Betrachtung widerspricht.

Da allen betrachteten Diskretisierungen die singuläre Diffusionsgleichung  $u_t = (\text{sgn } u_x)_x$  zugrundeliegt, untersuchen wir diese Gleichung im folgenden Abschnitt auf ihre unterschiedlichen Formulierungen und zugehörigen numerischen Eigenschaften.

### 3.4 Untersuchung der Gleichung $u_t = (\text{sgn } u_x)_x$

Die singuläre Diffusionsgleichung  $u_t = (\text{sgn } u_x)_x$  spielt nicht nur im Kontext unserer Filteralgorithmen für hyperbolische Erhaltungsgleichungen eine bedeutende Rolle, sondern ist auch in der Bildverarbeitung wohlbekannt. Da bei ihrer Anwendung die Totalvariation eines zugrundegelegten Bildes verringert wird, heißt diese Gleichung *TV flow*.

Im Folgenden vergleichen wir verschiedene Darstellungen und numerische Implementierungen dieser Gleichung und zeigen, daß einige Diskretisierungen zu neuen Oszillationen führen, so daß das resultierende Verfahren im Allgemeinen nicht TV-verringend ist. Dennoch können wir zeigen, daß *TV flow* selbst-stabilisierend wirkt, denn bei einer Verletzung der Stabilitätsbedingungen bleiben die resultierenden Oszillationen in ihrer Größe durch eine zur Zeitschrittweite proportionalen Konstante beschränkt.

### 3.4.1 Konsistente Kettenregel in 1D

Die Gleichung

$$\partial_t u = \nabla \cdot \left( \frac{\nabla u}{|\nabla u|} \right), \quad (3.37)$$

hat in den letzten Jahren großes Interesse im Bereich der Bildverarbeitung geweckt, wo sie unter den Namen *TV flow* oder *TV diffusion* bekannt ist. Die Filtergleichung zeichnet sich dadurch aus, daß sie keine Parameter benötigt, wohldefiniert ist [4, 7], die Form bestimmter Objekte erhält [7] und in endlicher Zeit zu konstanten Signalen führt. In [64] wurde gezeigt, daß (3.37) mit Anfangswerten  $f$  im Eindimensionalen äquivalent zur sogenannten *TV Regularisierung* [54] ist, bei der das Energiefunktional

$$E(u) := \int \left( (u - f)^2 + 2t|u_x| \right) dx$$

minimiert wird. Diese Variationsmethode ist bekannt für ihre guten Entrauschungseigenschaften.

Um die Probleme, die bei der Filtergleichung (3.37) auftreten können, aufzuzeigen, beschränken wir uns der Einfachheit halber auf den eindimensionalen Fall, für den Gleichung (3.37) durch

$$\partial_t u = \partial_x (\operatorname{sgn} \partial_x u) \quad (3.38)$$

gegeben ist. Selbst wenn  $\partial_x u$  beliebig glatt und damit differenzierbar ist, läßt sich die rechte Seite von Gleichung (3.38) nicht ohne weiteres mit Hilfe einer Kettenregel weder in  $BV$  noch im Distributionenraum  $\mathcal{D}'$  [72] berechnen. Daher beschäftigen wir uns zunächst mit der Herleitung einer konsistenten Kettenregel für die rechte Seite von Gleichung (3.38).

#### Direkte Berechnung

Wir betrachten Gleichung (3.38) und schreiben ihre rechte Seite gemäß

$$\operatorname{sgn} w(x) = (H \circ w)(x) - (H \circ -w)(x)$$

um, wobei  $H$  die Heaviside-Funktion bezeichnet. Sei  $x^*$  eine einfache Nullstelle von  $w$  im Intervall  $I^* := [a, b] := [x^* - \varepsilon, x^* + \varepsilon]$ ,  $\varepsilon > 0$ . In Anlehnung an [30] betrachten wir die Abbildung  $w$  als Koordinatenwechsel und müssen daher voraussetzen, daß  $w' \neq 0$  in der Umgebung von  $x^*$  gilt. Deshalb nehmen wir an, daß  $w'$  in  $I^*$  stets das gleiche Vorzeichen hat. Damit ergeben sich die beiden zu betrachtenden Fälle  $w'(x) > 0$  und  $w'(x) < 0$ .

**(i) Sei  $w'(x) > 0$  für alle  $x \in I^*$ :** Um  $H \circ (\pm w)$  als Funktion von  $x$  auszudrücken, sei darauf hingewiesen, daß die Differenz  $H(x - x^*) - H(x^* - x)$  gerade genau  $(H \circ w)(x) - (H \circ -w)(x)$  ist, was die Abbildungen 3.29 und 3.30 verdeutlichen.

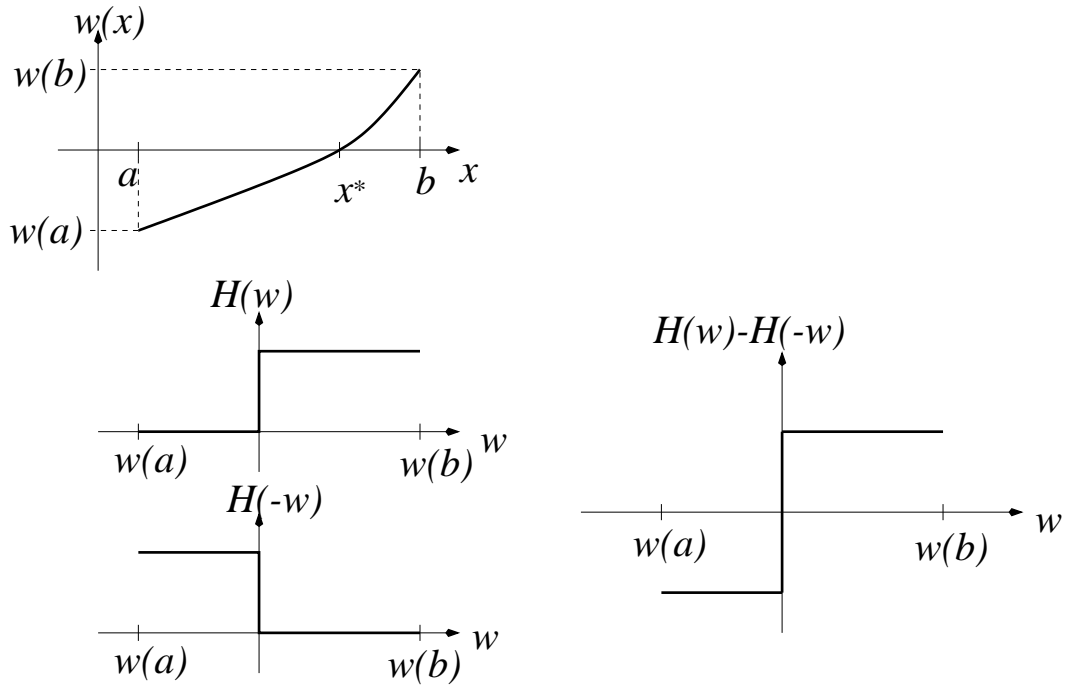
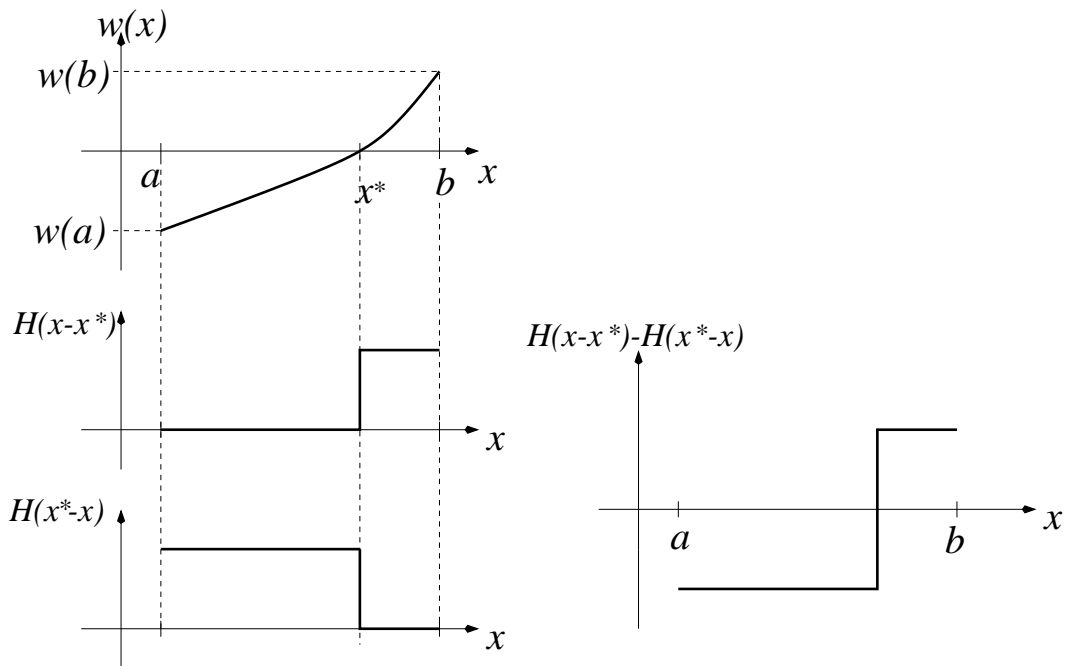
Damit ergibt sich

$$\operatorname{sgn} w(x) = H(x - x^*) - H(x^* - x),$$

und diese Gleichung kann im Sinn der Distributionen abgeleitet werden, was zu dem Ergebnis

$$\partial_x \operatorname{sgn} w(x) = 1 \cdot \delta(x - x^*) - (-1) \cdot \delta(x^* - x) = 2\delta(x - x^*) \quad (3.39)$$

führt.

Abbildung 3.29: Der Fall  $w' > 0$ Abbildung 3.30: Berechnungen für den Fall  $w' > 0$ 

(ii) Sei  $w'(x) < 0$  für alle  $x \in \mathbf{I}^*$ : Analoges Vorgehen wie im Fall (i) liefert

$$\begin{aligned} \partial_x \text{sgn } w(x) &= \partial_x (H(x^* - x) - H(x - x^*)) \\ &= (-1) \cdot \delta(x^* - x) - 1 \cdot \delta(x - x^*) = -2\delta(x - x^*). \end{aligned} \quad (3.40)$$

Fassen wir die Fälle (i) und (ii) zusammen, haben wir damit das folgende Lemma bewiesen:

**Lemma 3.6**

Sei  $w$  eine glatte Funktion mit einer einfachen Nullstelle  $x^*$  im Intervall  $[x^* - \varepsilon, x^* + \varepsilon]$ ,  $\varepsilon > 0$ , in dem  $w'(x) \neq 0$  gelte. Dann gilt

$$\partial_x \operatorname{sgn} w(x) = \operatorname{sgn} w'(x^*) \cdot 2\delta(x - x^*) \quad (3.41)$$

im Sinne der Distributionen.

### Kettenregel mit Hilfe von Regularisierungen

Eine andere Möglichkeit, die Kettenregel (3.41) zu erhalten, besteht in der Regularisierung der Signumfunktion durch den Ausdruck

$$\widetilde{\operatorname{sgn}}_m w(x) := \frac{2}{\pi} \arctan(mw(x)), \quad m \in \mathbb{N}. \quad (3.42)$$

Die zugehörige Ableitung ist durch

$$\partial_x \widetilde{\operatorname{sgn}}_m w(x) = \frac{2m}{\pi(1 + m^2 w(x)^2)} w'(x)$$

gegeben, wobei die Funktion  $\widetilde{\delta}_m(w(x)) := \frac{m}{\pi(1 + m^2 w(x)^2)}$  das Glied einer Delta-Folge ist, so daß

$$\widetilde{\delta}_m(w(x)) = \frac{m}{\pi(1 + m^2 w(x)^2)} \xrightarrow{\mathcal{D}'} \delta(w(x))$$

gilt und daher

$$\partial_x \widetilde{\operatorname{sgn}}_m w(x) \xrightarrow{\mathcal{D}'} 2\delta(w(x))w'(x)$$

folgt.

Setzen wir wiederum voraus, daß  $w$  eine isolierte Nullstelle in  $x^*$  hat und  $w' \neq 0$  in  $I^*$  gilt, folgt aufgrund der Transformationseigenschaften der Delta-Distribution,

$$\delta(w(x)) = \frac{\delta(x - x^*)}{|w'(x^*)|},$$

vergleiche [29] und [30], daß

$$\partial_x \widetilde{\operatorname{sgn}}_m w(x) \xrightarrow{\mathcal{D}'} 2\delta(w(x))w'(x) = \frac{w'(x)}{|w'(x^*)|} \cdot 2\delta(x - x^*)$$

gilt. Da  $\delta(x - x^*) \neq 0$  lediglich für  $x = x^*$  erfüllt ist, ist die letzte Aussage identisch mit

$$\partial_x \widetilde{\operatorname{sgn}}_m w(x) \xrightarrow{\mathcal{D}'} \frac{w'(x^*)}{|w'(x^*)|} \cdot 2\delta(x - x^*) = \operatorname{sgn} w'(x^*) \cdot 2\delta(x - x^*),$$

und dies ist gerade wieder (3.41).

### 3.4.2 Diskretisierungen und numerische Ergebnisse

Für die Konstruktion diskreter Filter, ausgehend von der kontinuierlichen Gleichung (3.38), haben wir nun drei mögliche Startpunkte, nämlich

1. Gleichung (3.38):  $\partial_t u = \partial_x (\text{sgn } \partial_x u)$ ,
2. Gleichung  $\partial_t u = 2 \text{sgn } (\partial_x^2 u(x^*)) \cdot \delta(x - x^*)$  unter Verwendung von (3.41) und
3. Gleichung  $\partial_t u = \partial_x (\widetilde{\text{sgn}}_m(\partial_x u))$  als regularisierte Version unter Verwendung von (3.42).

#### Direkte Diskretisierung

Wir beginnen mit der direkten Diskretisierung von Gleichung (3.38). Wie bei den vorangegangenen numerischen Anwendungen rechnen wir auf dem Gitter

$$\mathbb{G} := \{(x, t) \mid x = i\Delta x, t = n\Delta t, i \in \mathbb{Z}, n \in \mathbb{N}, \Delta x > 0, \Delta t > 0\}$$

und bezeichnen mit  $U_i^n$  die Approximation an  $u(i\Delta x, n\Delta t)$ . Darüber hinaus führen wir die Operatoren

$$\begin{aligned} E^\alpha U_i^n &:= U_{i+\alpha}^n, \quad \alpha \in \mathbb{Z}, \\ 1 U_i^n &:= E^0 U_i^n = U_i^n, \\ E U_i^n &:= E^1 U_i^n = U_{i+1}^n, \\ \Delta U_i^n &:= U_{i+1}^n - U_i^n = E U_i^n - U_i^n = (E - 1) U_i^n, \\ \nabla U_i^n &:= U_i^n - U_{i-1}^n = U_i^n - E^{-1} U_i^n = (1 - E^{-1}) U_i^n. \end{aligned}$$

ein. Der Einfachheit halber diskretisieren wir die zeitliche Ableitung stets durch die einfache Vorwärtsdifferenz

$$\partial_t u(i\Delta x, n\Delta t) = \frac{u_i^{n+1} - u_i^n}{\Delta t} + \mathcal{O}(\Delta t).$$

Mittels der Diskretisierung

$$\partial_x u(i\Delta x, n\Delta t) = \frac{u_{i+1}^n - u_{i-1}^n}{2\Delta x} + \mathcal{O}(\Delta x^2) = \frac{1}{2\Delta x} (E - E^{-1}) u_i^n + \mathcal{O}(\Delta x^2)$$

und dem Ersetzen der äußeren Ableitung in (3.38) durch die gleiche Differenz erhalten wir für die diskrete Version von (3.38) den Ausdruck

$$U_i^{n+1} = U_i^n + \frac{\Delta t}{2\Delta x} (\text{sgn } (U_{i+2}^n - U_i^n) - \text{sgn } (U_i^n - U_{i-2}^n)).$$

Dieses diskrete Analogon leidet offensichtlich an einer Entkopplung von geraden und ungeraden Gitterpunkten, die durch eine Reduktion der räumlichen Schrittweite behoben werden kann. Wir erhalten

$$U_i^{n+1} = U_i^n + \frac{\Delta t}{\Delta x} (\text{sgn } (U_{i+1}^n - U_i^n) - \text{sgn } (U_i^n - U_{i-1}^n)), \quad (3.43)$$

also gerade den bereits bekannten Signumfilter (3.31). Im Gegensatz zu bisherigen Betrachtungen ersetzen wir nun die Signumfunktion durch die Differenz zweier Heaviside-Funktionen und formulieren die inneren Differenzen mit Hilfe der eingeführten Operatoren. Dieses Vorgehen liefert

$$\begin{aligned}
U_i^{n+1} &= U_i^n + \frac{\Delta t}{\Delta x} ([H((E-1)U_i^n - H((1-E)U_i^n] \\
&\quad - [H((1-E^{-1})U_i^n) - H((E^{-1}-1)U_i^n)]) \\
&= U_i^n + \frac{\Delta t}{\Delta x} (H \circ (E-1) - H \circ (1-E) \\
&\quad - H \circ (1-E^{-1}) + H \circ (E^{-1}-1)) U_i^n \\
&= [1 + \frac{\Delta t}{\Delta x} (H \circ (E-1) - H \circ (1-E) \\
&\quad - H \circ (1-E^{-1}) + H \circ (E^{-1}-1))] U_i^n.
\end{aligned} \tag{3.44}$$

Mit Hilfe numerischer Experimente untersuchen wir im Folgenden die Eigenschaften des Operators

$$\begin{aligned}
C(\Delta t) := \\
[1 + \frac{\Delta t}{\Delta x} (H \circ (E-1) - H \circ (1-E) - H \circ (1-E^{-1}) + H \circ (E^{-1}-1))] .
\end{aligned} \tag{3.45}$$

Als zu filterndes Anfangssignal verwenden wir, wie schon bei anderen Anwendungen, die numerische Lösung des instabilen Zentrale-Differenzen-Schemas zweiter Ordnung für das Anfangswertproblem (1.49), das in Abbildung 3.31 dargestellt ist. Das zugrundeliegende räumliche Gitter besteht aus 400 äquidistanten Gitterpunkten. Im Gegensatz zur Filterimplementie-

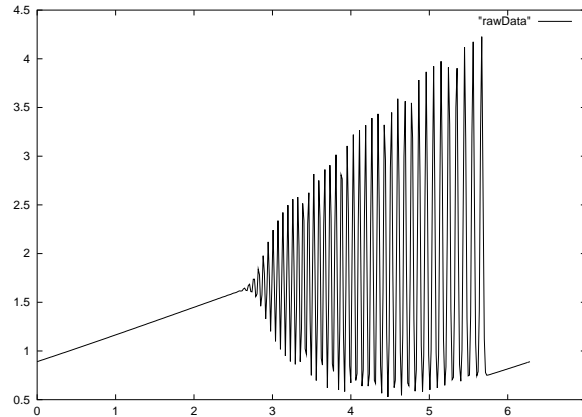


Abbildung 3.31: Zu filterndes Signal

rung nach jedem Zeitschritt des Differenzenschemas wenden wir die verschiedenen diskreten Filteralgorithmen jetzt erst nach Beendigung des Differenzenverfahrens, also auf die fertige Differenzenlösung, an.

Zunächst wählen wir den Gitterkoeffizienten  $\lambda := \frac{\Delta t}{\Delta x} = 0.0005$  und führen nacheinander 10000, 20000 und 40000 Iterationen von Filtergleichung (3.44) aus. Die numerischen Ergebnisse sind in den Abbildungen 3.32 und 3.33 dargestellt. Wie man klar erkennt, verringert der Filter (3.44) die Totalvariation des Anfangssignals deutlich und schrumpft darüber hinaus

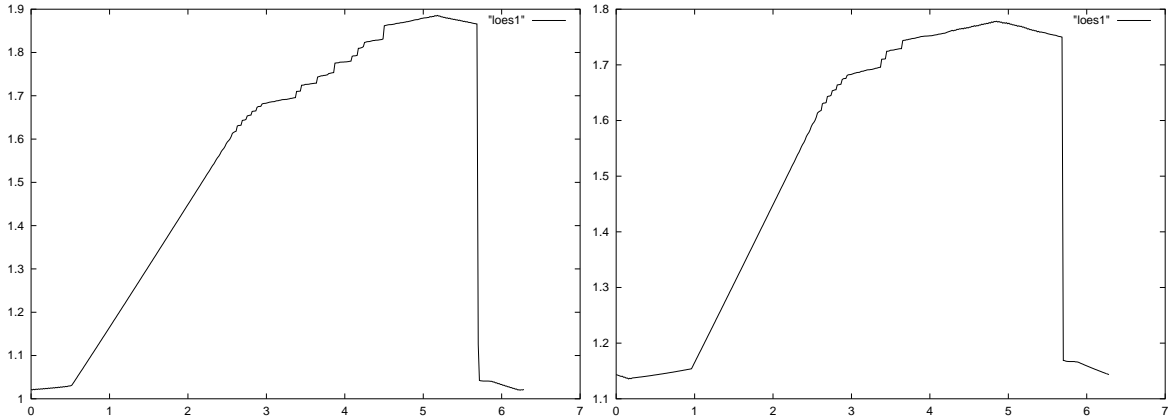


Abbildung 3.32: Gefiltertes Signal nach 10000 (links) und 20000 Iterationen

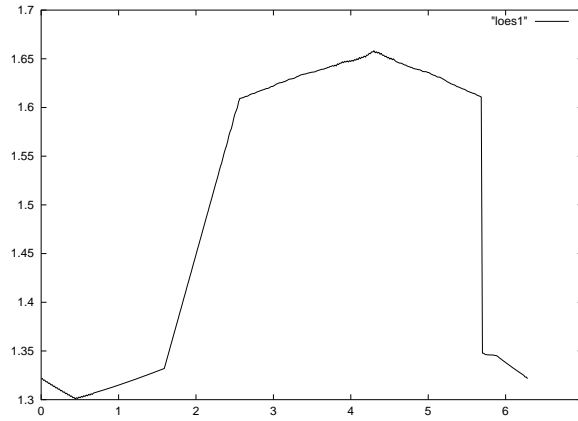


Abbildung 3.33: Gefiltertes Signal nach 40000 Iterationen

das gesamte Signal mit zunehmender Anzahl von Zeitschritten immer stärker. Dennoch sind in den Lösungen selbst nach 40000 Zeitschritten noch hochfrequente Oszillationen zu beobachten. Die Darstellung der Totalvariation in Abhängigkeit von der Anzahl der Zeitschritte, dargestellt im linken Teil von Abbildung 3.34, zeigt, daß die Totalvariation, zum Zeitniveau  $n$  berechnet durch  $\sum_{i=1}^{399} |U_{i+1}^n - U_i^n|$ , verringert wird. Allerdings ergibt die genaue Betrachtung der Totalvariation zwischen den Zeitschritten 39900 und 40000, daß tatsächlich noch Oszillationen vorhanden sind, wie Abbildung 3.34 zeigt. Vergrößern wir den Gitterkoeffizienten auf  $\lambda = 0.5$ , erhalten wir nach 10000 Filteriterationen die in Abbildung 3.35 dargestellte Lösung mit entsprechender Totalvariation. Das gefilterte Signal ist überall verrauscht und die Totalvariation wächst auf ein bestimmtes mittleres Niveau an, um das herum sich beschränkte Oszillationen bilden. Zu beachten ist, daß die hochfrequenten Oszillationen im gefilterten Signal anscheinend von der Größenordnung  $\lambda$  sind. Eine nähere Betrachtung der Totalvariation für die ersten 100 Iterationen sowie für die Zeitschritte zwischen 7000 und 7010 in Abbildung 3.36 verdeutlicht die allgemeine Struktur der Totalvariation. Weitere numerische Experimente zeigen, daß der Gitterkoeffizient  $\lambda$  ohne Einschränkung weiter vergrößert werden kann, was einen deutlichen Unterschied zu der strengen CFL-Bedingung, die bei expliziten Zeitdiskretisierungen zu erwarten ist, darstellt. Abbildung 3.37 zeigt beispielsweise das gefilterte Signal und die zugehörige Totalvariation für  $\lambda = 5$ . Das resultierende Signal weist starke

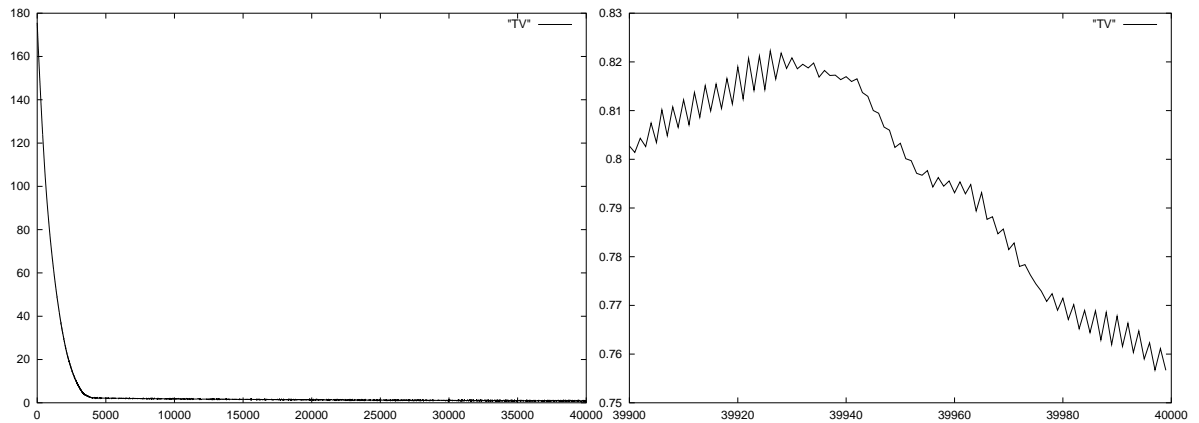


Abbildung 3.34: Totalvariation in Abhängigkeit der Zeitschritte

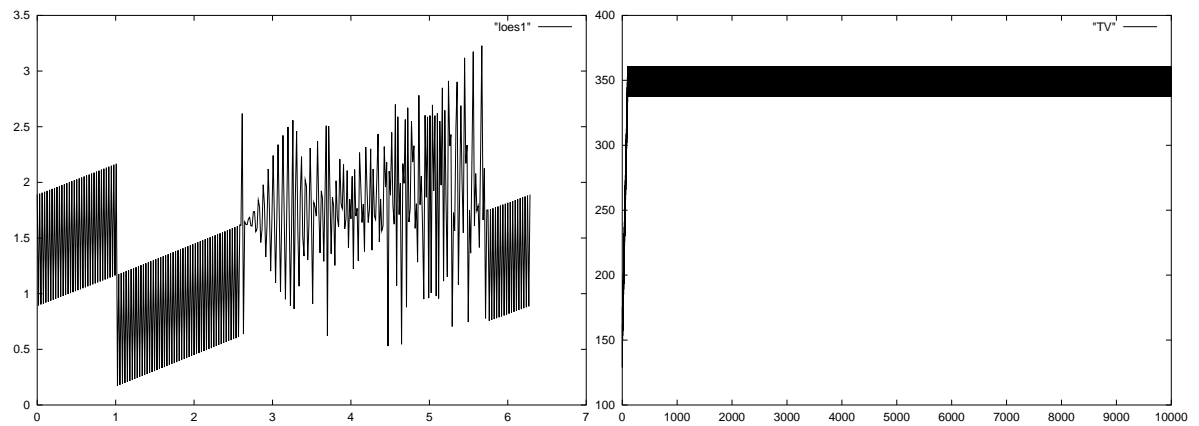


Abbildung 3.35: Gefiltertes Signal (links) und Totalvariation

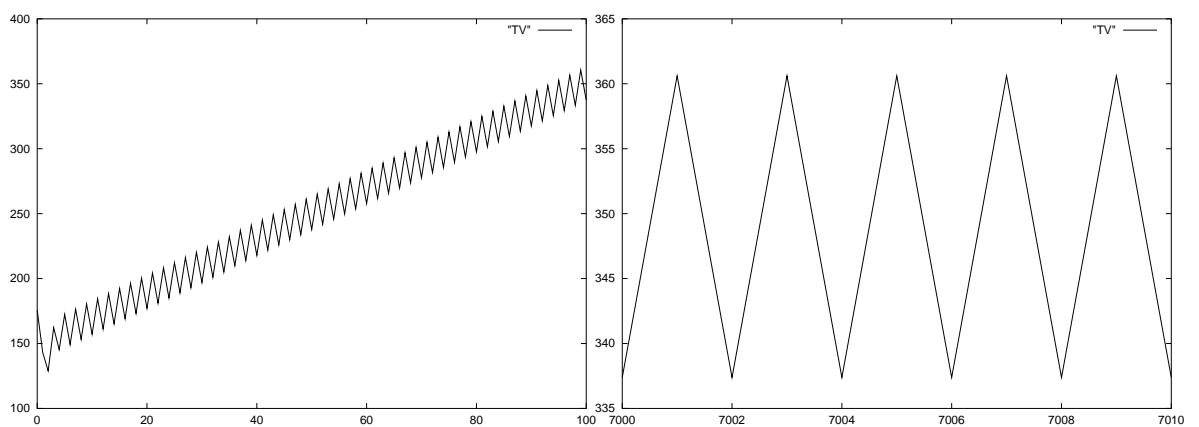


Abbildung 3.36: Totalvariation in den ersten 100 Zeitschritten (links) und zwischen den Zeitschritten 7000 und 7010

Oszillationen auf, und die Totalvariation ist weitaus größer als für das ungefilterte Signal.



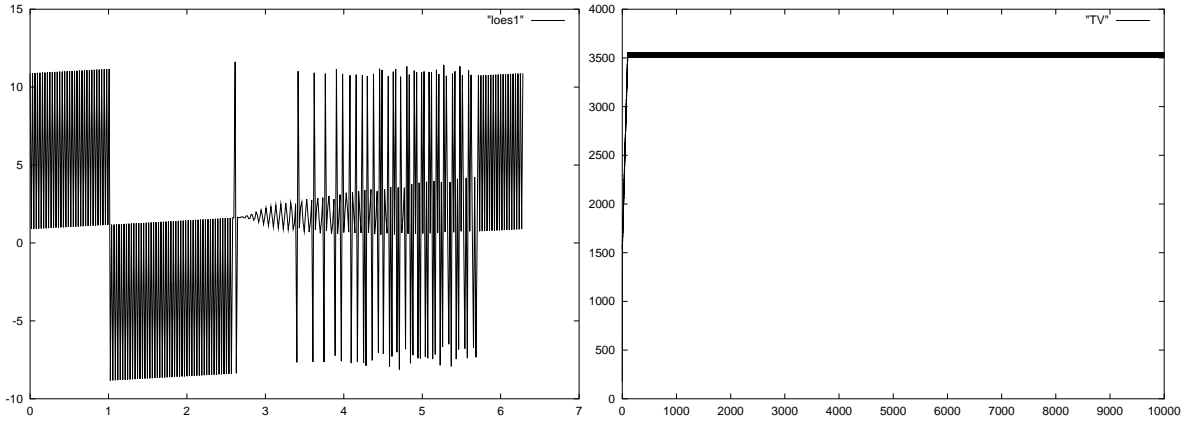


Abbildung 3.37: Gefiltertes Signal (links) und Totalvariation

**Bemerkung 3.7**

Die numerischen Verfahren (3.43) und (3.44) sind vollständig analog, da die Signumfunktion exakt durch die Differenz zweier Heaviside-Funktionen dargestellt wird.

**Theorem 3.8**

Der in (3.45) definierte diskrete Operator  $C(\Delta t)$  erzeugt auf Gittergröße im Wechsel Minima und Maxima und verursacht Oszillationen der Amplitude  $2\lambda = 2\Delta t/\Delta x$  auf folgende Weise: Wenn  $\max\{|U_i^n - U_{i-1}^n|, |U_i^n - U_{i+1}^n|\} < \frac{\Delta t}{\Delta x}$  gilt, entsteht ein neuer Extremwert an der Stelle  $x_i$ .

**Beweis** Sei an der Stelle  $x_i$  ein Maximum, so daß  $U_i^n > U_{i-1}^n$  und  $U_i^n > U_{i+1}^n$  gelte. Dann gilt aufgrund von (3.43)

$$\begin{aligned} U_i^{n+1} &= U_i^n + \frac{\Delta t}{\Delta x} \left( \underbrace{\text{sgn}(U_{i+1}^n - U_i^n)}_{=-1} - \underbrace{\text{sgn}(U_i^n - U_{i-1}^n)}_{=1} \right) \\ &= U_i^n - 2 \frac{\Delta t}{\Delta x} = U_i^n - 2\lambda. \end{aligned}$$

Für  $\max\{|U_i^n - U_{i-1}^n|, |U_i^n - U_{i+1}^n|\} < 2\frac{\Delta t}{\Delta x}$  wird damit aus dem Maximum ein Minimum. Analoges ergibt sich für den Fall eines Minimums an der Stelle  $x_i$ . ■

**Korollar 3.9**

Der in (3.45) definierte diskrete Operator  $C(\Delta t)$  ist im Gegensatz zu seinem kontinuierlichen Modell, aus dem er entwickelt wurde, sowohl  $TV$ -erhöhend als auch  $TV$ -verringend, je nach Wahl des Gitterkoeffizienten  $\lambda$ . Dennoch ist die Totalvariation für den stationären Fall beschränkt durch die Größenordnung von  $\lambda$ . Selbst für ein explizites Zeitschrittverfahren mit großer Zeitschrittweite bleiben die Oszillationen beschränkt. Mittels (3.44) ergibt sich

$$TV(U^{n+1}) := \sum_i |U_{i+1}^{n+1} - U_i^{n+1}| \leq \left(1 + \frac{\Delta t}{\Delta x} \|\mathcal{C}\|\right) TV(U^n)$$

mit  $\mathcal{C} := H \circ (E - 1) - H \circ (1 - E) - H \circ (1 - E^{-1}) + H \circ (E^{-1} - 1)$  und  $0 \leq \|\mathcal{C}\| \leq 2$ .

Abhilfe für das Problem der nicht abnehmenden Totalvariation läßt sich schaffen, indem wir die diskrete Gleichung (3.44) mit kleinem  $\lambda$  so lange anwenden, bis die durch  $\lambda$  bestimmte untere Grenze der Totalvariation erreicht ist. Dann wechseln wir zu einem weitaus einfacheren Filter, beispielsweise der linearen Wärmeleitungsgleichung, um mit Hilfe dieser Gleichung die restlichen Oszillationen zu glätten. Bemerkenswert ist, daß dieses Vorgehen einem neuen, von Steidel und Mitarbeitern [64] entwickelten Filteralgorithmus ähnelt, der aus der genauen Betrachtung eines Zwei-Pixel-Modells resultiert.

Bei der vorgeschlagenen Problemlösung wird das Stabilitätsproblem auf Kosten der Konsistenz des Verfahrens gelöst. Doch da Oszillationen gebildet werden, wenn der Gradient nahe bei Null ist, ist das Einfügen einer linearen Diffusion an dieser Stelle in der Praxis unproblematisch.

Eine andere, oft verwendete Möglichkeit, die Oszillationen in der Lösung zu vermeiden, ist das Regularisieren der *TV flow*-Gleichung durch den Ausdruck

$$\partial_t = \partial_x \left( \frac{\partial_x u}{\sqrt{\beta^2 + (\partial_x u)^2}} \right), \quad (3.46)$$

wobei  $\beta > 0$  ein betragsmäßig kleiner Regularisierungsparameter ist. Die Diffusivität ist damit beschränkt durch  $\frac{1}{\beta}$  und für

$$\frac{\Delta t}{\Delta x^2} \leq \frac{1}{2\beta}$$

läßt sich leicht zeigen, daß die direkte Diskretisierung von Gleichung (3.46) ein diskretes Maximum-Minimum-Prinzip erfüllt und damit stabil ist.

Für

$$\frac{\Delta t}{\Delta x^2} \leq \frac{1}{4\beta}$$

gilt sogar Vorzeichenstabilität, vergleiche [18], so daß der Filter in diesem Fall TV-vermindernd ist.

Dennoch zeigen die angeführten Stabilitätsschranken, daß wir bei Annäherung an das *TV flow*-Modell, also für den Grenzwert  $\beta \rightarrow 0$ , infinitesimal kleine Zeitschrittweiten für unser explizites Verfahren erhalten.

### Diskretisierung mit Hilfe der Kettenregel

Als nächstes entwickeln wir eine Diskretisierung von (3.38) mit Hilfe der Kettenregel (3.41) und untersuchen ihre numerischen Eigenschaften.

Für eine isolierte Nullstelle  $x^*$  von  $\partial_x u$  ist die kontinuierliche Filtergleichung durch

$$\partial_t u = 2 \operatorname{sgn}(\partial_x^2 u(x^*)) \cdot \delta(x - x^*)$$

gegeben. Als Indikator für Nullstellen von  $\partial_x u$  verwenden wir das Produkt  $\Delta U_i^n \cdot \nabla U_i^n$ , das im Fall einer Extremalstelle negativ ist. Mit Hilfe der Heaviside-Funktion erhalten wir damit

$$H(-\Delta U_i^n \cdot \nabla U_i^n) = \begin{cases} 1 & ; \text{ Nullstelle bei } i \text{ in der diskreten Ableitung} \\ 0 & ; \text{ keine Nullstelle bei } i \end{cases}$$

als Indikator. Die zweite Ableitung diskretisieren wir gemäß

$$\begin{aligned}\partial_x^2 u(x^*)|_{x=i\Delta x} &= H(-\Delta U_i^n \cdot \nabla U_i^n) \cdot \left( \frac{U_{i+1}^n - 2U_i^n + U_{i-1}^n}{\Delta x^2} + \mathcal{O}(\Delta x^2) \right) \\ &= H(-\Delta U_i^n \cdot \nabla U_i^n) \cdot \left( \frac{(E - 2 + E^{-1})U_i^n}{\Delta x^2} + \mathcal{O}(\Delta x^2) \right),\end{aligned}$$

und die Dirac-Funktion an einer Nullstelle  $x_i^*$  der diskreten Ableitung von  $u$  wird approximiert durch

$$\tilde{\delta}(x - x_i^*) := \begin{cases} \frac{1}{\Delta x} & ; \quad x_i^* - \frac{\Delta x}{2} < x_i^* < x_i^* + \frac{\Delta x}{2} \\ 0 & ; \quad \text{sonst} \end{cases},$$

wodurch eine zulässige Dirac-Folge mit  $\int \tilde{\delta}(x - x_i^*) dx = 1$  definiert wird. Ersetzen wir die Signum-Funktion wieder durch die Differenz zweier Heaviside-Funktionen, ergibt sich insgesamt

$$\begin{aligned}U_i^{n+1} &= U_i^n + 2\frac{\Delta t}{\Delta x} H(-\Delta U_i^n \cdot \nabla U_i^n) \cdot (H \circ (E - 2 + E^{-1}) \\ &\quad - H \circ (-(E - 2 + E^{-1}))) U_i^n \\ &= \left( 1 + 2\frac{\Delta t}{\Delta x} H(-\Delta U_i^n \cdot \nabla U_i^n) \cdot (H \circ (E - 2 + E^{-1}) - H \circ (-(E - 2 + E^{-1}))) \right) U_i^n\end{aligned}$$

als neues Finite-Differenzen-Verfahren.

Die numerischen Ergebnisse sind, wie erwartet, denen der direkten Diskretisierung sehr ähnlich. Abbildung 3.38 zeigt das gefilterte Signal nach 10000 Zeitschritten mit  $\lambda = 0.5$  sowie die zugehörige Totalvariation. Wie im Fall der direkten Diskretisierung führt die Diskretisierung mit Hilfe der Kettenregel zu Oszillationen in der Lösung, die die Totalvariation erhöhen können.

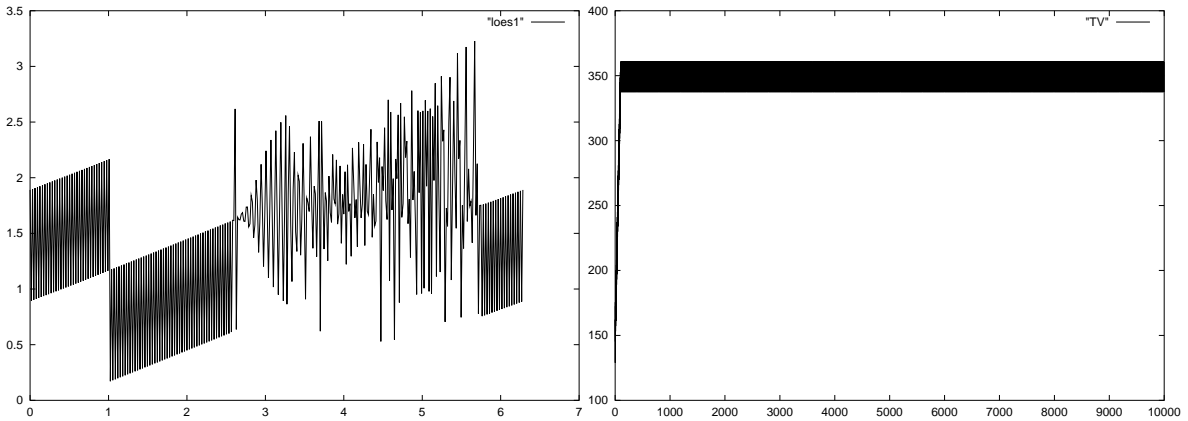


Abbildung 3.38: Gefiltertes Signal (links) und Totalvariation

## Diskretisierung mittels Regularisierung

Als letztes betrachten wir die regularisierte Gleichung

$$\partial_t u = \partial_x (\widetilde{\text{sgn}}_m(\partial_x u)) = \frac{2}{\pi} \partial_x (\arctan(m \partial_x u)). \quad (3.47)$$

Eine zugehörige Diskretisierung erhalten wir, indem wir den Zentrale-Differenzen-Operator  $(E - E^{-1})$  sowohl für die äußere als auch für die innere Ableitung anwenden und anschließend die Weite der einbezogenen Gitterpunkte so verringern, daß keine Entkopplung zwischen geraden und ungeraden Gitterpunkten mehr auftreten kann. Es ergibt sich

$$U_i^{n+1} = U_i^n + \frac{2\Delta t}{\pi\Delta x} \left( \arctan(m(U_{i+1}^n - U_i^n)) - \arctan(m(U_i^n - U_{i-1}^n)) \right). \quad (3.48)$$

Schon an dieser Stelle läßt sich erkennen, daß die numerischen Lösungen dieser Differenzengleichung nicht sehr anders als die des vorher betrachteten Verfahrens sein werden. Da  $|\arctan(x)| < \frac{\pi}{2}$  gilt, ergibt die Differenz der Arctan-Funktion in (3.48) im ungünstigsten Fall  $\pm\pi$  und wir erhalten

$$U_i^{n+1} = U_i^n \pm \frac{2\Delta t}{\Delta x} = U_i^n \pm 2\lambda,$$

so daß in der numerischen Lösung wieder Oszillationen der Größenordnung  $2\lambda$  auftreten. Tatsächlich weisen Berechnungen mit  $\lambda = 0.05$  und  $m = 1$  ein vielversprechendes Verhalten auf, da das gefilterte Signal sehr glatt erscheint, wie Abbildung 3.39 zeigt.

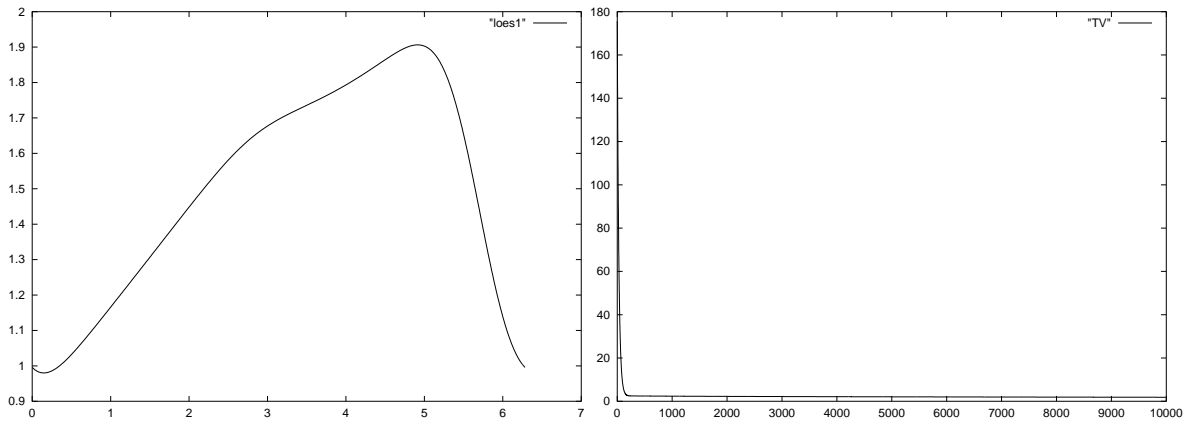


Abbildung 3.39: Gefiltertes Signal (links) und Totalvariation

Erhöhen wir den Regularisierungsparameter  $m$  auf  $m = 100$  und behalten  $\lambda = 0.05$  bei, ergibt sich die in Abbildung 3.40 dargestellte Lösung. Offensichtlich wird die diskrete Gleichung (3.48) für wachsendes  $m$  in die direkte Diskretisierung (3.44) überführt, und Oszillationen von der Größenordnung  $\lambda$  treten auf. Dies läßt sich deutlich erkennen, wenn wir  $m = 100$  konstant lassen und  $\lambda$  auf  $\lambda = 5$  erhöhen, was in Abbildung 3.41 dargestellt ist. Die numerische Lösung ist nun vollständig durch Oszillationen der Amplitude  $\lambda$  zerstört.

Die Regularisierung führt damit nur zu einem sinnvollen Filteralgorithmus, wenn der Regularisierungsparameter  $m$  klein ist, was einer starken Regularisierung entspricht.

Das gleiche Ergebnis ergibt sich, wenn wir erst die Ableitung in (3.48) ausführen und anschließend diskretisieren. Dann erhalten wir eine Regularisierung der Delta-Distribution, die ebenfalls nur für kleine Regularisierungsparameter gute Resultate liefert.

Damit haben wir drei verschiedene diskrete Modelle der singulären Diffusionsgleichung (3.38) vorgestellt und gezeigt, daß alle drei Versionen Oszillationen in der numerischen Lösung erzeugen können, so daß die Totalvariation der Lösung zunimmt, obwohl die kontinuierliche Diffusionsgleichung TV-vermindernd ist. Damit haben wir einen Widerspruch zwischen den

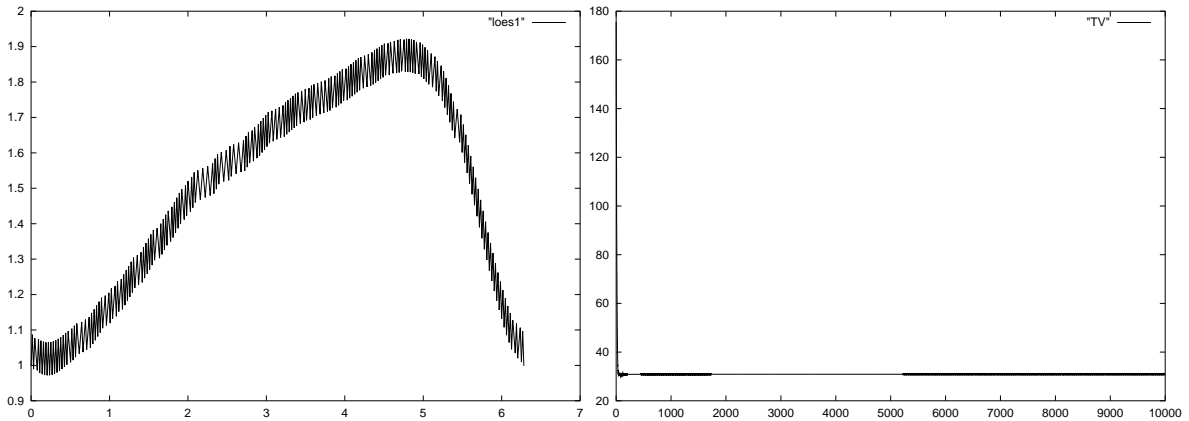


Abbildung 3.40: Gefiltertes Signal (links) und Totalvariation

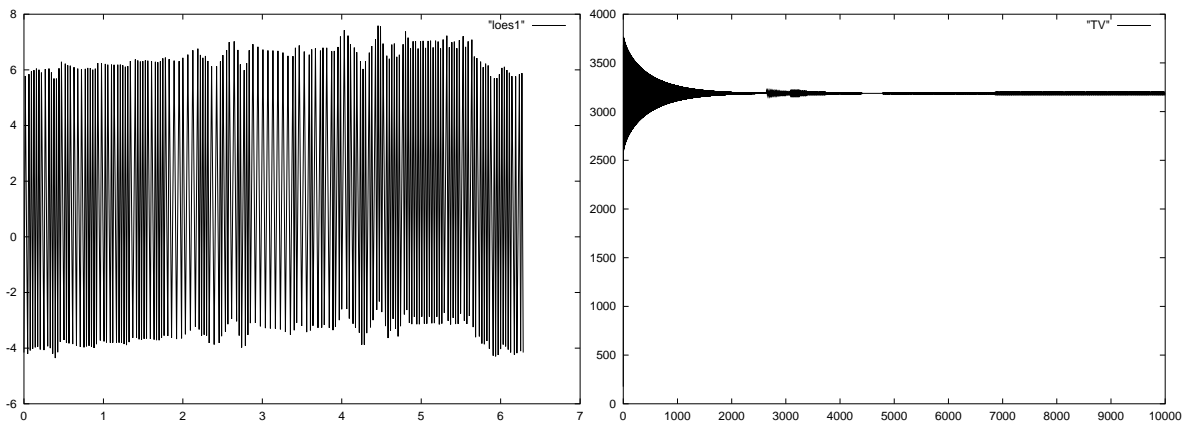


Abbildung 3.41: Gefiltertes Signal (links) und Totalvariation

Eigenschaften einer kontinuierlichen Gleichung und denen des zugehörigen diskreten Modells aufgezeigt.

Naheliegend ist der Versuch, die Diskretisierung mittels Regularisierung nicht mit einem expliziten sondern einem impliziten Zeitschrittverfahren zu diskretisieren, da die Anwendung eines expliziten Zeitschrittverfahrens bei singulären Diffusionsgleichungen fragwürdig sein kann. Allerdings zeigen die Untersuchungen in [9], daß auch die implizite Diskretisierung gleiche Probleme mit sich bringt.



## 4 Entwicklung neuer Filtermethoden

Im vorangegangenen Kapitel haben wir nichtlineare diskrete Filter, entwickelt in der Bildverarbeitung, zum Lösen hyperbolischer Erhaltungsgleichungen angewendet, verglichen und mit Hilfe neuer Diskretisierungen modifiziert.

Im Folgenden orientieren wir uns zunächst an den vorhandenen Filteralgorithmen von Engquist, Lötstedt und Sjögreen und entwickeln auf dieser Grundlage ein neues kontinuierliches Filtermodell, das wir auf seine mathematischen und numerischen Eigenschaften hin untersuchen.

Da sich unter anderem ein sehr einfacher Filter ergibt, der, ebenso wie der TV-Filter von Chan, Osher und Shen, eine Verschiebung der Unstetigkeit erzeugt, untersuchen wir sein Amplituden- und Phasenverhalten am Ende dieses Abschnittes.

Im zweiten Teil des Kapitels wird ein neuer Ansatz zur Implementierung diskreter Filteralgorithmen vorgestellt und für die neu entwickelte Filtergleichung ausgeführt. Mit Hilfe einer Diffusionsanalyse werden die zugehörigen Diffusionen ermittelt und direkt in das zugrundegelegte Differenzenverfahren integriert, so daß wir eine weitere neue Filtervariante für das Lösen hyperbolischer Erhaltungsgleichungen erhalten.

Im letzten Teil des Kapitels entwickeln wir, gelöst von jeglicher kontinuierlicher Modellvorstellung, neue vollständig diskrete Filteralgorithmen, die wir für skalare Probleme und Systeme von Erhaltungsgleichungen implementieren.

### 4.1 Extrema-Diffusion

Zunächst betrachten wir noch einmal den diskreten Filter von Engquist, Lötstedt und Sjögreen, vergleichen seine Wirkungsweise mit verschiedenen Modellen der Bildverarbeitung und entwickeln aus den verschiedenen Ansätzen eine neue Filterroutine, deren Eigenschaften wir mathematisch und numerisch untersuchen.

### 4.1.1 Motivation und Herleitung des kontinuierlichen Filtermodells

In Kapitel 2 haben wir den diskreten Filter (2.31) von Engquist, Lötstedt und Sjögreen in seine zugehörige kontinuierliche Form umgewandelt und diese rein formal mit dem TV-vermindernden Filtermodell

$$\partial_t u = -|\nabla u|F(L(u)) \quad (4.1)$$

der Bildverarbeitung verglichen. Vergleichen wir dagegen die Wirkungsweise des diskreten Filters (2.31) mit den Eigenschaften des TV-Filters (4.1), stellen wir in der anschaulichen Funktionalität der Filter große Unterschiede fest. Der TV-Filter bewirkt das Aufsteilen von geglätteten Kanten, denn für  $u_{xx} \neq 0$  liefert er zusätzliche positive oder negative Diffusion, je nachdem, ob  $u_{xx}$  größer oder kleiner Null ist, was zur Kantenverschärfung führt. Für  $u_{xx} = 0$ , also an Stellen ohne gekrümmte Kanten, liefert der Filter den Wert Null, so daß sich die Lösung nicht ändert. Der Filter von Engquist, Lötstedt und Sjögreen wird dagegen an Extremwerten, also dort, wo  $u_x = 0$  gilt, aktiv. An diesen Stellen glättet der Filter die einzelnen Extremwerte durch einseitige Diffusion mit einem der Nachbarwerte. Da dabei der Reduktionsbetrag des Extremums gleich dem Korrekturbetrag des Nachbarpunktes ist und stets ein Betrag addiert und der andere subtrahiert wird, ist die Filterroutine konservativ. Weil die Konservativität eine grundlegende Filtereigenschaft für das numerische Lösen hyperbolischer Erhaltungsgleichungen ist, wollen wir sie bei der Konstruktion neuer Filterroutinen beibehalten und nehmen daher Abstand von dem nichtkonservativen TV-Modell  $\partial_t u = -|\nabla u|F(L(u))$ . Als Vorbild für eine kontinuierliche, konservative Filtergleichung betrachten wir stattdessen die *TV flow*-Gleichung

$$\partial_t u = \nabla \cdot \left( \frac{\nabla u}{|\nabla u|} \right)$$

aus der Bildverarbeitung, die aufgrund ihrer Divergenzform konservativ ist. Ziel ist es, ein kontinuierliches Filtermodell zu konstruieren, das ebenfalls in Divergenzform formuliert werden kann, und dessen anschauliches Verhalten mit dem diskreten Filter von Engquist, Lötstedt und Sjögreen übereinstimmt.

Eine neue Filterroutine, die der letzten Forderung, der Diffusion von Extremwerten, gerecht wird, kann durch

$$\partial_t u = (1 - \operatorname{sgn}(|\nabla u|)) \operatorname{div} \left( \frac{\partial u}{\partial \eta} \right)$$

dargestellt werden, wobei die Divergenz der Richtungsableitung gebildet wird und  $\eta = \frac{\nabla u}{|\nabla u|}$  gilt. Daraus ergibt sich das Filtermodell

$$\partial_t u = (1 - \operatorname{sgn}(|\nabla u|)) \Delta u, \quad (4.2)$$

das zwar noch nicht konservativ ist, aber die gewünschte Extrema-Diffusion umsetzt. Mit Hilfe der Divergenzform ergibt sich eine konservative Version dieses Filters gemäß

$$\partial_t u = \operatorname{div} ((1 - \operatorname{sgn}(|\nabla u|)) \nabla u). \quad (4.3)$$

Der Faktor  $(1 - \operatorname{sgn}(|\nabla u|))$  liefert für Extrema den Wert eins und in jedem anderen Fall Null, wodurch gewährleistet wird, daß die neue Filterroutine, wie bei Engquist, Lötstedt



und Sjögren, nur an Extremwerten aktiviert wird. Die Größe des Korrekturterms wird von  $\operatorname{div}(\nabla u)$  beziehungsweise  $\Delta u$  bestimmt. Beide Filtermodelle sind so konstruiert, daß bei ihrer Anwendung keine neuen Extremwerte im Datensatz entstehen können.

### Zusammenhang zum Perona–Malik–Modell

Bemerkenswert ist, daß sich Verbindungen zwischen den beiden hergeleiteten Filtermodellen (4.2), (4.3) und dem bekannten Perona–Malik–Modell (2.17), das wir in Kapitel 2.1.2 beschrieben haben, aufzeigen lassen.

Perona und Malik [47] entwickelten 1987 erstmals kontinuierliche Formulierungen für adaptive Glättungsmethoden der Bildverarbeitung mit Hilfe nichtlinearer partieller Differentialgleichungen. Um die Diffusion adaptiv zu steuern, konstruierten sie den nichtlinearen Diffusionsfilter (2.17),

$$\partial_t u = \operatorname{div}(c(|\nabla u|^2)\nabla u). \quad (4.4)$$

Als klassische Diffusivität verwenden sie unter anderem den Ausdruck (2.20), der sich auch in der Form

$$c(s^2) = \frac{1}{1 + s^2/\lambda^2}, \quad \lambda > 0,$$

darstellen läßt, wobei  $\lambda$  ein Kontrastparameter ist, der zwischen Regionen, die hohe beziehungsweise niedrige Diffusion erfordern, unterscheidet.

Im eindimensionalen Fall ist (4.4) durch

$$\partial_t u = [c(u_x^2(t, x))u_x(t, x)]_x \quad (4.5)$$

gegeben, und mit der Flußfunktion  $\Phi(s) := s \cdot c(s^2)$  läßt sich Gleichung (4.5) darstellen als

$$\partial_t u = \Phi'(u_x)u_{xx}. \quad (4.6)$$

Die Flußfunktion erfüllt dabei  $\Phi'(s) \geq 0$  für  $|s| \leq \lambda$  und  $\Phi'(s) < 0$  für  $|s| > \lambda$ , wie auch in Abbildung 4.1 zu erkennen ist.

Der Parameter  $\lambda$  teilt die Diffusionsgleichung (4.5) in zwei Fälle auf, für  $|u_x| \leq \lambda$  ergibt sich eine Differentialgleichung vom vorwärts parabolischen Typ, für  $|u_x| > \lambda$  eine vom rückwärts parabolischen Typ.

Betrachten wir nach diesen Ausführungen wieder unser neu konstruiertes konservatives Filtermodell (4.3) in einer Raumdimension, also

$$\partial_t u = ((1 - \operatorname{sgn}|u_x|)u_x)_x, \quad (4.7)$$

lassen sich gewisse Ähnlichkeiten zum Perona–Malik–Modell feststellen. Die Funktion  $c$  findet sich in (4.7) in dem Ausdruck  $(1 - \operatorname{sgn}|u_x|)$  wieder, allerdings mit dem Unterschied, daß  $c$  in unserem Modell nicht mehr von  $|u_x|^2$  sondern von  $|u_x|$  abhängt. Anschaulich läßt sich die Diffusivität  $(1 - \operatorname{sgn}|u_x|)$  als Grenzfall der Perona–Malik–Diffusivität auffassen. Die von ihnen gewählte Funktion  $c$  hängt von  $\lambda$  ab und konvergiert für  $\lambda \rightarrow 0$  gegen eine Funktion, die

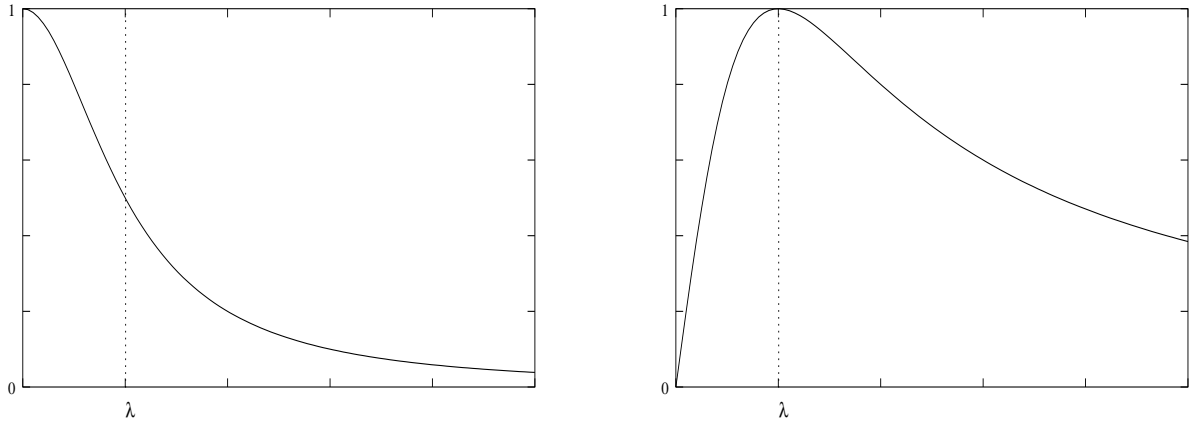


Abbildung 4.1: Diffusivität  $c(s) = \frac{1}{1+s^2/\lambda^2}$  (links) und Flußfunktion  $\Phi(s) = \frac{s}{1+s^2/\lambda^2}$  (rechts)

für  $|u_x| = 0$  den Wert eins hat und sonst Null ist. Dieser Grenzwert ist gerade die Diffusivität  $c(|u_x|) = 1 - \operatorname{sgn} |u_x|$ .

Motiviert durch den Filter von Engquist, Lötstedt und Sjögreen und dem *TV flow* der Bildverarbeitung haben wir damit einen neuen Filter entwickelt, der sich als Spezial- beziehungsweise Grenzfall des Perona–Malik–Modells interpretieren läßt.

In den folgenden Abschnitten wird das neue Filtermodell auf seine mathematischen und numerischen Eigenschaften hin untersucht.

### 4.1.2 Äquivalenz der konservativen und nichtkonservativen Modellgleichung

Zunächst beschäftigen wir uns mit den beiden verschiedenen Formulierungen des kontinuierlichen Filtermodells und zeigen, daß die konservative und die nichtkonservative Filterdarstellung mathematisch übereinstimmen.

Zur Vereinfachung beschränken wir uns auf den eindimensionalen Fall, für den die konservative Filtergleichung durch

$$u_t = ((1 - \operatorname{sgn} |u_x|)u_x)_x \quad (4.8)$$

gegeben ist. Das zugehörige nichtkonservative Filtermodell besitzt in 1D die Darstellung

$$u_t = (1 - \operatorname{sgn} |u_x|)u_{xx}. \quad (4.9)$$

#### Lemma 4.1

Sei  $u_x$  eine glatte Funktion mit einer einfachen Nullstelle  $x^*$  im Intervall  $[x^* - \varepsilon, x^* + \varepsilon]$ ,  $\varepsilon > 0$ , in dem  $u_{xx} \neq 0$  gelte. Dann sind die Filtergleichungen (4.8) und (4.9) im Sinne der Distributionen äquivalent.

**Beweis** Die konservative Formulierung (4.8) läßt sich zunächst zu

$$\begin{aligned} u_t &= ((1 - \operatorname{sgn} |u_x|)u_x)_x \\ &= (1 - \operatorname{sgn} |u_x|)u_{xx} + u_x(-\operatorname{sgn} |u_x|)_x \\ &= (1 - \operatorname{sgn} |u_x|)u_{xx} - u_x(\operatorname{sgn} |u_x|)_x \end{aligned} \quad (4.10)$$

umformen. Den Term  $(\operatorname{sgn} |u_x|)_x$  berechnen wir, ähnlich wie in Abschnitt 3.4.1, mit Hilfe der Heaviside-Funktion. Sei dazu  $w$  mit  $w(x) := u_x$  eine glatte Funktion mit einer einfachen Nullstelle  $x^*$  im Intervall  $I^* = [x^* - \varepsilon, x^* + \varepsilon]$ ,  $\varepsilon > 0$ , in dem  $w'(x) \neq 0$  gelte. Daher nehmen wir an, daß  $w'$  in  $I^*$  stets das gleiche Vorzeichen besitzt, und wir haben die Fälle  $w'(x) > 0$  und  $w'(x) < 0$  zu unterscheiden.

Sei daher zunächst  $w'(x) > 0$ . Dann gilt im Sinne der Distributionen

$$\operatorname{sgn} |u_x| = \operatorname{sgn} |w(x)| = H(|w(x)|) = H(w(x)) + H(-w(x))$$

und damit

$$\begin{aligned} \frac{d}{dx} \operatorname{sgn} |w(x)| &= \frac{d}{dx} (H(w(x)) + H(-w(x))) \\ &= \frac{d}{dx} (H(x - x^*) + H(x^* - x)) = \delta(x - x^*) + (-\delta(x^* - x)) \\ &= \delta(x - x^*) - \delta(-(x - x^*)) = \delta(x - x^*) - \delta(x - x^*) = 0. \end{aligned}$$

Das gleiche Ergebnis läßt sich mit Hilfe der Testfunktionen erzielen. Sei dazu  $\phi$  eine beliebige Testfunktion mit kompaktem Träger. Dann gilt

$$\begin{aligned} \left\langle \frac{d}{dx} \operatorname{sgn} |w(x)|, \phi \right\rangle &= -\langle \operatorname{sgn} |w(x)|, \phi' \rangle \\ &= -\int_{\mathbb{R}} (H(x - x^*) + H(x^* - x)) \phi'(x) dx \\ &= -\int_{\mathbb{R}} H(x - x^*) \phi'(x) dx - \int_{\mathbb{R}} H(x^* - x) \phi'(x) dx \\ &= -\int_{x^*}^{\infty} \phi'(x) dx - \int_{-\infty}^{x^*} \phi'(x) dx \\ &= \phi(x^*) - \phi(x^*) = 0. \end{aligned}$$

Für den Fall  $w'(x) < 0$  ergibt sich durch analoge Berechnung ebenfalls  $\frac{d}{dx} \operatorname{sgn} |w(x)| = 0$ , so daß wir insgesamt

$$(\operatorname{sgn} |u_x|)_x = 0$$

erhalten.

Damit ergibt sich für Gleichung (4.10)

$$u_t = (1 - \operatorname{sgn} |u_x|)u_{xx} - u_x(\operatorname{sgn} |u_x|)_x = (1 - \operatorname{sgn} |u_x|)u_{xx},$$

und die Äquivalenz von konservativer und nichtkonservativer kontinuierlicher Filterdarstellung ist gezeigt. ■

Da das konservative und das nichtkonservative Filtermodell mathematisch übereinstimmen, diskretisieren wir in den folgenden Abschnitten beide Formulierungen und prüfen, ob auch ihr numerisches Verhalten vergleichbar ist. Dabei gehen wir sowohl auf die direkten Diskretisierungen der beiden Filtergleichungen als auch auf eine technisch ausgefeiltere Diskretisierung ein.

### 4.1.3 Direkte Diskretisierung der konservativen Filtergleichung

Wir beginnen mit der direkten Diskretisierung des eindimensionalen konservativen Filters

$$u_t = ((1 - \operatorname{sgn} |u_x|)u_x)_x, \quad (4.11)$$

indem wir mit Vorwärts- und Rückwärtsdifferenzen so diskretisieren, daß jeweils die Werte der direkten Nachbarn bei der Berechnung eines neuen Punktwertes berücksichtigt werden. Es ergibt sich

$$u_t = \frac{(1 - \operatorname{sgn} |u_x|)u_x|_{i+1} - (1 - \operatorname{sgn} |u_x|)u_x|_i}{\Delta x}$$

und damit

$$U_i^{n+1} = U_i^n + \frac{\Delta t}{\Delta x^2} (d_{i+1}(U_{i+1}^n - U_i^n) - d_i(U_i^n - U_{i-1}^n)),$$

wobei  $d_i$  die Diskretisierung von  $(1 - \operatorname{sgn} |u_x|)|_i$  ist und

$$d_i = \begin{cases} 1 & , \text{ falls ein Extremwert in } i \text{ existiert} \\ 0 & , \text{ sonst} \end{cases}$$

gilt.

Für die numerische Anwendung sind die Filterterme aus Konsistenzgründen zusätzlich mit  $\Delta x$  zu multiplizieren, so daß die Filtergleichung die Gestalt

$$U_i^{n+1} = U_i^n + \frac{\Delta t}{\Delta x} (d_{i+1}(U_{i+1}^n - U_i^n) - d_i(U_i^n - U_{i-1}^n)) \quad (4.12)$$

annimmt.

## Numerische Resultate

Wir implementieren die Filterroutine (4.12) in das Cauchy–Problem (1.49) für Burgers’ Gleichung, gelöst mit einem Zentrale–Differenzen–Verfahren zweiter Ordnung. Ohne Filteranwendung ergibt sich bei einer CFL–Zahl von 0.009 auf einem Gitter aus 400 äquidistanten Gitterpunkten nach  $t = 2.5$  die in Abbildung 3.9 dargestellte oszillierende Lösung. Integrieren wir nach jedem Zeitschritt des Differenzenverfahren eine Iteration der Filtergleichung (4.12), erhalten wir dagegen die im linken Teil von Abbildung 4.2 dargestellte Lösung. Die Oszillationen sind stark geglättet, die Steilheit und Lage des Stoßes sind erhalten, wie ein Vergleich mit der zugehörigen Lax–Friedrichs–Lösung zeigt, und die CFL–Zahl konnte auf 0.09 erhöht werden, was eine Beschleunigung des Differenzenverfahrens bewirkt.

Betrachten wir allerdings das Lösungsverhalten für  $\Delta x \rightarrow 0$ , stellen wir fest, daß die Lösung nicht konvergiert, wie aus dem rechten Teil von Abbildung 4.2 ersichtlich wird, in dem die gleiche Lösung wie links dargestellt ist, jetzt allerdings mit 1600 Gitterpunkten statt der ursprünglichen 400.

Weitere numerische Experimente zeigen, daß die Oszillationen in der Lösung nicht nur für kleiner werdende Gitterweite sondern auch bei steigender Anzahl von Filteriterationen pro

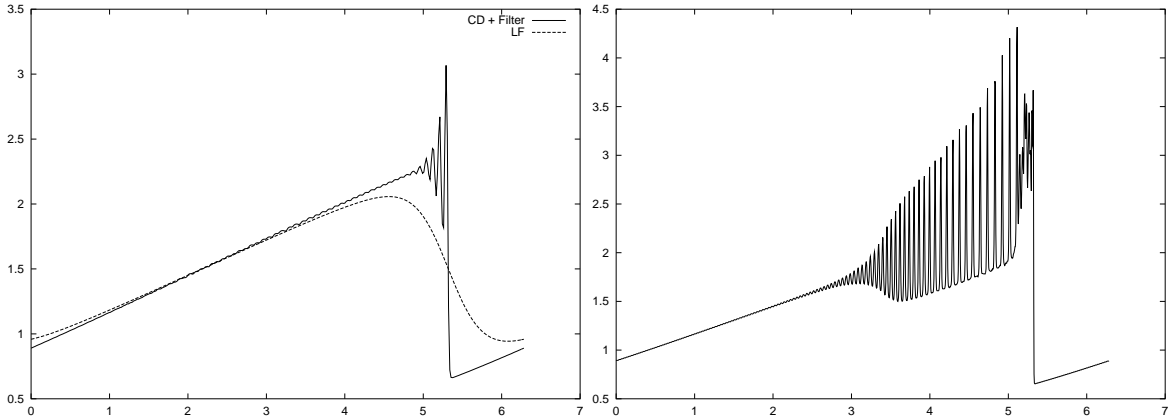


Abbildung 4.2: Gefilterte Lösung im Vergleich zur Lax-Friedrichs-Lösung für  $N = 400$  (links) und  $N = 1600$  (rechts)

Zeitschritt zunehmen. Dieses Verhalten ändert sich auch nicht bei weiterer Multiplikation der Filterterme mit  $\Delta x$  oder bei Verkleinerung der Zeitschrittweite  $\Delta t$ . Ersetzen wir das instabile Zentrale-Differenzen-Verfahren durch das Lax-Wendroff-Verfahren und wiederholen die Simulationen, stellen wir fest, daß die Oszillationen auch in diesem Fall mit steigender Anzahl von Filteriterationen pro Zeitschritt zunehmen. Damit ist es offensichtlich, daß das Entstehen der Oszillationen durch die Filterdiskretisierung (4.12) und nicht durch das Zentrale-Differenzen-Verfahren verursacht wird.

Alle weiteren direkten Diskretisierungen von (4.11), die wir mit Hilfe von einseitigen Vorwärts- und Rückwärtsdifferenzen oder zentralen Diskretisierungen erzielen können, führen zu gleich schlechten numerischen Resultaten.

## Stabilitätsanalyse

Um die Problematik der direkt diskretisierten Filtergleichung auch mathematisch belegen zu können, führen wir im Folgenden eine Stabilitätsanalyse nach von Neumann [26, 3] für die Diskretisierung (4.12) durch.

Die allgemeine Herleitung der von Neumannschen Stabilitätsanalyse ist in Kapitel 4.1.6 beschrieben, so daß wir an dieser Stelle auf die Herleitung verzichten und direkt mit der Anwendung des Verfahrens auf die Filtergleichung (4.12) beginnen.

Um Verwechslungen mit der komplexen Zahl  $i$  zu vermeiden, bezeichnen wir die diskreten Gitterpunkte während der Stabilitätsanalyse mit  $(x_j, t_n)$ ,  $j = \dots, -1, 0, 1, \dots$ , die Approximation von  $u(x_j, t_n)$  mit  $U_j^n$  und die Diffusionsfaktoren mit  $d_j$ . Der Ansatz von von Neumann ist durch

$$U_j^n := e^{at} e^{ik_m x}$$

gegeben, wobei  $k_m \in \mathbb{R}$  die Wellenzahl bezeichnet und  $a \in \mathbb{C}$  gilt. Die Bedingung für Stabilität lautet

$$\left| \frac{U_j^{n+1}}{U_j^n} \right| = |e^{a\Delta t}| =: |G| \leq 1 \quad \text{für alle } \beta = k_m \Delta x.$$

Um die diskrete Gleichung (4.12) auf Stabilität zu untersuchen, sind aufgrund der verschiedenen Werte von  $d_j$  und  $d_{j+1}$  drei Fälle zu unterscheiden:

**1. Fall**  $d_{j+1} = d_j = 1$ :

$$U_j^{n+1} = U_j^n + \frac{\Delta t}{\Delta x} (U_{j+1}^n - 2U_j^n + U_{j-1}^n)$$

und mit dem Ansatz von von Neumann ergibt sich

$$\begin{aligned} e^{a(t+\Delta t)} e^{ik_m x} &= e^{at} e^{ik_m x} + \frac{\Delta t}{\Delta x} \left( e^{at} e^{ik_m(x+\Delta x)} - 2e^{at} e^{ik_m x} + e^{at} e^{ik_m(x-\Delta x)} \right) \\ \Leftrightarrow e^{a\Delta t} &= 1 + \frac{\Delta t}{\Delta x} \left( e^{ik_m \Delta x} - 2 + e^{-ik_m \Delta x} \right) \\ &= 1 + \frac{\Delta t}{\Delta x} (2 \cos(k_m \Delta x) - 2) \\ &= 1 + \lambda (2 \cos \beta - 2), \quad \lambda = \frac{\Delta t}{\Delta x}, \\ &= 1 - 2\lambda + 2\lambda \cos \beta \\ \Rightarrow |e^{a\Delta t}| &= |1 - 2\lambda + 2\lambda \cos \beta| \stackrel{!}{\leq} 1, \end{aligned}$$

und die Stabilitätsbedingung  $|G| \leq 1$  ist für diesen Fall für  $0 \leq \lambda \leq 1/2$  erfüllt.

**2. Fall**  $d_{j+1} = 1, d_j = 0$ :

$$U_j^{n+1} = U_j^n + \frac{\Delta t}{\Delta x} (U_{j+1}^n - U_j^n)$$

und analoge Berechnungen führen zu

$$\begin{aligned} e^{a\Delta t} &= 1 + \frac{\Delta t}{\Delta x} (\cos \beta + i \sin \beta - 1) \\ &= 1 - \lambda + \lambda (\cos \beta + i \sin \beta) \\ \Rightarrow |e^{a\Delta t}| &\leq |1 - \lambda| + \lambda |\cos \beta + i \sin \beta| \\ &= |1 - \lambda| + \lambda \sqrt{\cos^2 \beta + \sin^2 \beta} \\ &= |1 - \lambda| + \lambda \stackrel{!}{\leq} 1, \end{aligned}$$

so daß die Bedingung  $|G| \leq 1$  für  $0 \leq \lambda \leq 1$  erfüllt ist.

**3. Fall**  $d_{j+1} = 0, d_j = 1$ :

$$U_j^{n+1} = U_j^n - \frac{\Delta t}{\Delta x} (U_j^n - U_{j-1}^n)$$

und es ergibt sich

$$\begin{aligned} e^{a\Delta t} &= 1 - \lambda + \lambda (\cos \beta - i \sin \beta) \\ \Rightarrow |e^{a\Delta t}| &\leq |1 - \lambda| + \lambda |\cos \beta - i \sin \beta| \\ &= |1 - \lambda| + \lambda, \end{aligned}$$

so daß das Verfahren, wie im zweiten Fall, für  $0 \leq \lambda \leq 1$  stabil ist.

Damit ist die Stabilitätsbedingung für alle möglichen Varianten der diskreten Gleichung (4.12) erfüllt, wenn  $\lambda \leq 1/2$  gilt, so daß Instabilitäten nicht der Grund für das fehlerhafte numerische Verhalten der Filtergleichung sind.

### Modifizierte Gleichungen

Ein weiteres Problem bei der numerischen Approximation kontinuierlicher Gleichungen ist die Tatsache, daß numerische Verfahren zwar konvergieren, aber nicht mehr notwendigerweise gegen die gesuchte, wahre Lösung der Problemstellung.

Wir untersuchen unser numerisches Verfahren auf diese Fehlerquelle hin, indem wir die verschiedenen diskreten Fälle von (4.12) mit Hilfe der Taylorentwicklung wieder auf kontinuierliche Form bringen und überprüfen, ob dies noch der zugehörigen Filtergleichung entspricht. Wir betrachten wieder die obigen Fälle der Filtergleichung für die verschiedenen Werte von  $d_{i+1}$  und  $d_i$  im einzelnen.

**1. Fall**  $d_{i+1} = d_i = 1$ :

$$U_i^{n+1} = U_i^n + \frac{\Delta t}{\Delta x}(U_{i+1}^n - 2U_i^n + U_{i-1}^n)$$

Mit Hilfe der Taylorentwicklung nach der Variablen  $x$  ergibt sich aus dieser diskreten Gleichung die kontinuierliche Form

$$u_t = hu_{xx} + \frac{1}{2}h^3u_{xxxx} + O(h)^5 = hu_{xx} + O(h)^3,$$

wobei  $h$  die räumliche Schrittweite  $\Delta x$  bezeichnet. Da wir bei der Implementierung der diskreten Filtergleichung aus Konsistenzgründen mit  $\Delta x$  multipliziert haben, tritt dieser Faktor natürlich auch in der kontinuierlichen Gleichung auf, so daß diese Formulierung genau der Ausgangsgleichung

$$u_t = ((1 - \operatorname{sgn} |u_x|)u_x)_x$$

entspricht, wenn man zusätzlich berücksichtigt, daß der Vorfaktor  $(1 - \operatorname{sgn} |u_x|)$  auf den Wert 1 gesetzt wurde. Damit approximiert der diskrete Filter in diesem Fall die richtige Gleichung.

**2. Fall**  $d_{i+1} = 1, d_i = 0$ :

$$U_i^{n+1} = U_i^n + \frac{\Delta t}{\Delta x}(U_{i+1}^n - U_i^n)$$

Für diese Gleichung liefert die Taylorentwicklung

$$u_t = u_x + \frac{h}{2}u_{xx} + \frac{h^2}{6}u_{xxx} + O(h)^3.$$

Es zeigt sich, daß der zusätzliche Term  $u_x$  in der kontinuierlichen Formulierung auftritt, so daß in diesem Fall nicht mehr die tatsächliche Filtergleichung approximiert wird, wenn  $u_x \neq 0$  gilt. Tritt diese Kombination von  $d_i$  und  $d_{i+1}$  bei der numerischen Anwendung wirklich auf und gilt  $u_x \neq 0$ , kann dieser Term das fehlerhafte numerische Verhalten des Filters begründen. Ähnliches ergibt sich für den

**3. Fall**  $d_{i+1} = 0, d_i = 1$ :

$$U_i^{n+1} = U_i^n - \frac{\Delta t}{\Delta x}(U_i^n - U_{i-1}^n).$$

Mit Taylorentwicklung ergibt sich

$$u_t = -u_x + \frac{h}{2}u_{xx} - \frac{h^2}{6}u_{xxx} + O(h)^3,$$

und auch hier tritt der zusätzliche Term  $-u_x$  auf, so daß auch in diesem Fall nicht die richtige kontinuierliche Gleichung approximiert wird, wenn  $u_x \neq 0$  gilt.

Überprüfen wir bei der numerischen Anwendung der Filtergleichung (4.12), ob die Fälle 2 und 3 tatsächlich auftreten, stellen wir fest, daß dies der Fall ist. Das fehlerhafte numerische Verhalten der diskreten Filtergleichung (4.12) kann daher darauf beruhen, daß für verschiedene Fälle die falsche kontinuierliche Gleichung approximiert wird. Allerdings bleibt die Frage offen, ob  $u_x$  in diesen Fällen tatsächlich von Null verschieden ist, oder ob das Zusammenspiel der verschiedenen Filterfälle nicht doch die richtige kontinuierliche Gleichung approximiert, was die Tatsache, daß für Fall 2 der Term  $u_x$  und für Fall 3 der Term  $-u_x$  auftritt, vermuten läßt.

Für alle anderen möglichen direkten Diskretisierungen von (4.11) ergeben sich die gleichen numerischen und analytischen Eigenschaften, so daß deutlich wird, daß technisch ausgefeiltere Methoden für die Diskretisierung der Modellgleichung (4.11) notwendig sind.

Zuvor betrachten wir noch die direkte Diskretisierung der nichtkonservativen Filtergleichung

$$u_t = (1 - \operatorname{sgn}|u_x|)u_{xx},$$

um zu überprüfen, ob hier die gleichen Diskretisierungsprobleme auftreten.

#### 4.1.4 Direkte Diskretisierung der nichtkonservativen Filtergleichung

Die nichtkonservative Formulierung der kontinuierlichen Filtergleichung ist durch

$$u_t = (1 - \operatorname{sgn}|u_x|)u_{xx} \tag{4.13}$$

gegeben. Der Vorfaktor  $(1 - \operatorname{sgn}|u_x|)$  wird wie zuvor an der Stelle  $i$  durch  $d_i$  approximiert, wobei  $d_i$  gemäß

$$d_i = \begin{cases} 1 & , \text{ falls ein Extremwert in } i \text{ existiert} \\ 0 & , \text{ sonst} \end{cases}$$

berechnet wird. Für  $u_{xx}$  wählen wir eine zentrale Diskretisierung und erhalten die diskrete Darstellung

$$U_i^{n+1} = U_i^n + \frac{\Delta t}{\Delta x^2} d_i (U_{i+1}^n - 2U_i^n + U_{i-1}^n).$$

der kontinuierlichen Gleichung (4.13).

Für die numerische Implementierung des Filters sind die Filterterme aus Konsistenzgründen zusätzlich mit  $\Delta x$  zu multiplizieren, so daß der diskrete Filter insgesamt in der Form

$$U_i^{n+1} = U_i^n + \frac{\Delta t}{\Delta x} d_i (U_{i+1}^n - 2U_i^n + U_{i-1}^n) \tag{4.14}$$

gegeben ist.



## Numerische Resultate

Wir implementieren diese Filterroutine in das Cauchy-Problem (1.49) für Burgers' Gleichung, gelöst mit dem Zentrale-Differenzen-Verfahren zweiter Ordnung. Da der Filter zur Stabilisierung des zentralen Verfahrens notwendig ist, integrieren wir eine Iteration der Gleichung (4.14) nach jedem Zeitschritt des Differenzenverfahrens. Auf einem äquidistanten Gitter aus 800 Gitterpunkten ergibt sich bei einer CFL-Zahl von 0.1 nach  $t = 2.5$  die im linken Teil von Abbildung 4.3 dargestellte Lösung.

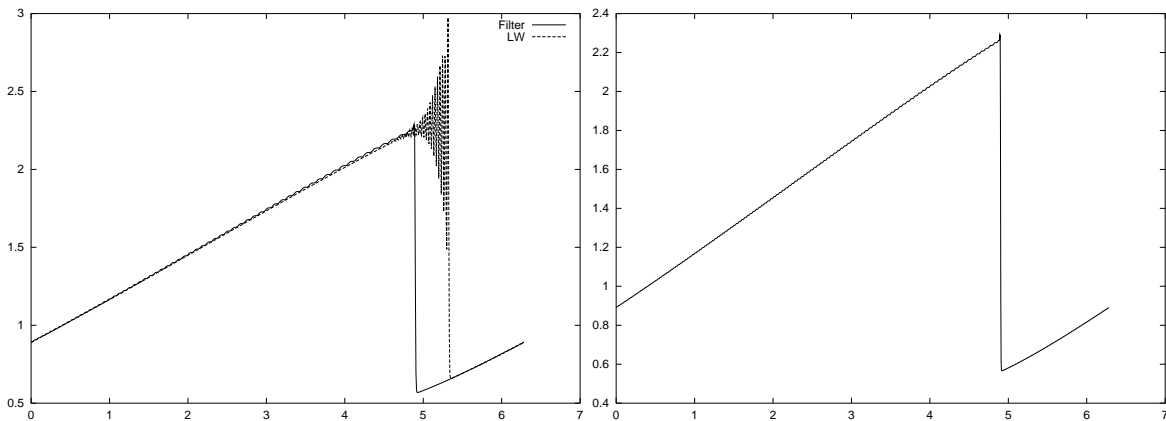


Abbildung 4.3: Gefilterte Lösung im Vergleich zur Lax-Wendroff-Lösung für  $N = 800$  (links) und gefilterte Lösung für  $N = 1600$  (rechts)

Zwar sind im linearen Anteil der Lösung noch kleine Oszillationen zu erkennen, die durch das instabile Zentrale-Differenzen-Schema verursacht werden, aber für  $\Delta x \rightarrow 0$  verschwinden diese Störungen und im Gegensatz zum konservativen Filter konvergiert das Verfahren, wie im rechten Teil der Abbildung 4.3 zu erkennen ist. Auch das Problem, daß mit steigender Anzahl von Filteriterationen pro Zeitschritt neue Oszillationen entstehen, tritt bei der nichtkonservativen Filterroutine nicht auf. Allerdings leidet die Lösung an einem deutlichen Verschiebungsfehler, wie ein Vergleich mit der zugehörigen Lax-Wendroff-Lösung zeigt. Er läßt sich darauf zurückführen, daß die implementierte Filtergleichung nicht mehr konservativ ist und die Datenpunkte vor und im Stoß daher fast ausschließlich nach unten korrigiert werden.

Insgesamt liefert der diskrete Filter (4.14) zwar ein sehr stabiles und konvergentes numerisches Verfahren, das Oszillationen fast vollständig glättet und gleichzeitig die Steilheit und Höhe des Stoßes erhält, aber aufgrund fehlender Konservativität einen Verschiebungsfehler in der Lösung verursacht.

Ein Vergleich mit den Ergebnissen des konservativen Filters (4.12) zeigt, wie unterschiedlich die numerischen Eigenschaften der beiden diskreten Filter sind, obwohl ihre zugehörigen kontinuierlichen Gleichungen mathematisch äquivalent sind. Damit wird deutlich, daß sich Eigenschaften kontinuierlicher Filtergleichungen nicht unbedingt auf die daraus entwickelten diskreten Filter übertragen lassen.

### 4.1.5 Konservative Diskretisierung der konservativen Filtergleichung

Wir betrachten das konservative Filtermodell

$$u_t = ((1 - \operatorname{sgn} |u_x|)u_x)_x, \quad (4.15)$$

für das wir im Folgenden eine konservative Diskretisierung entwickeln.

Die Diskretisierung des Faktors  $(1 - \operatorname{sgn} |u_x|)$  an der Stelle  $i$  durch  $d_i$  mit

$$d_i = \begin{cases} 1 & , \text{ falls ein Extremwert in } i \text{ existiert} \\ 0 & , \text{ sonst} \end{cases}$$

behalten wir bei. Das kontinuierliche Modell ist damit durch

$$u_t = (du_x)_x$$

gegeben, und wir diskretisieren wie folgt:

$$\begin{aligned} \frac{U_i^{n+1} - U_i^n}{\Delta t} &= (du_x)_x|_i \\ &= \frac{du_x|_{i+1} - du_x|_i}{\Delta x} \\ &= \frac{1}{\Delta x} \left( \frac{d_{i+1} + d_i}{2} \frac{U_{i+1}^n - U_i^n}{\Delta x} - \frac{d_{i-1} + d_i}{2} \frac{U_i^n - U_{i-1}^n}{\Delta x} \right) \\ &= \frac{1}{2\Delta x^2} ((d_{i+1} + d_i)(U_{i+1}^n - U_i^n) + (d_{i-1} + d_i)(U_i^n - U_{i-1}^n)). \end{aligned}$$

Bei der Implementierung dieser diskreten Gleichung wird die rechte Seite aus Konsistenzgründen zusätzlich mit  $\Delta x$  multipliziert, so daß die neue Filterroutine die Darstellung

$$U_i^{n+1} = U_i^n + \frac{\Delta t}{2\Delta x} ((d_{i+1} + d_i)(U_{i+1}^n - U_i^n) + (d_{i-1} + d_i)(U_i^n - U_{i-1}^n)) \quad (4.16)$$

erhält.

### Filterimplementierung für eindimensionale skalare Probleme

Wir beginnen die numerischen Testfälle mit dem 1D Cauchy-Problem (1.49) für Burgers' Gleichung, das wir zunächst mit Hilfe des Lax-Wendroff-Verfahrens

$$\begin{aligned} U_i^{n+1} = U_i^n &- \frac{\lambda}{2}(f(U_{i+1}^n) - f(U_{i-1}^n)) \\ &+ \frac{\lambda^2}{2} \left[ \frac{1}{2}(U_i^n + U_{i+1}^n)(f(U_{i+1}^n) - f(U_i^n)) - \frac{1}{2}(U_{i-1}^n + U_i^n)(f(U_i^n) - f(U_{i-1}^n)) \right] \end{aligned}$$

zweiter Ordnung mit  $f(u) = \frac{1}{2}u^2$  approximativ lösen. Auf einem Gitter aus 400 äquidistanten Gitterpunkten ergibt sich bei einer CFL-Zahl von 0.1 nach  $t = 2.5$  die im linken Teil von

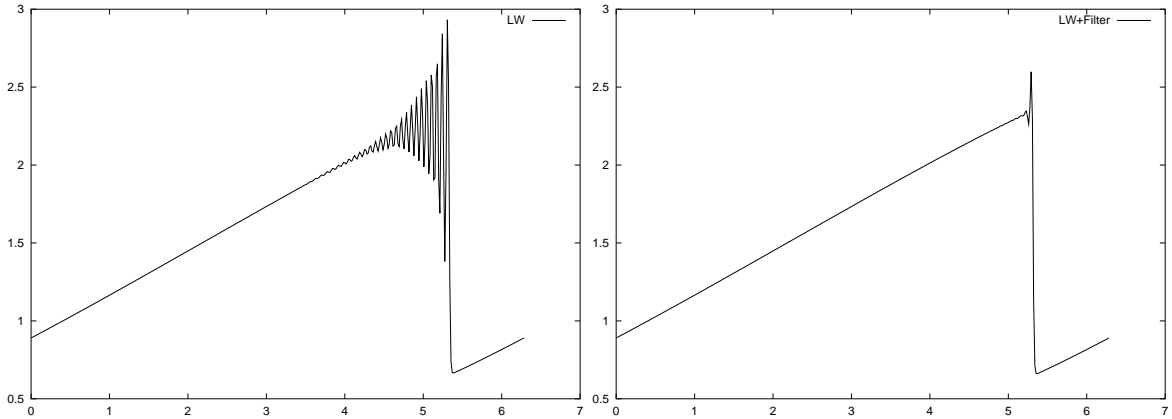


Abbildung 4.4: Lax-Wendroff-Lösung (links) und nach jedem Zeitschritt gefilterte Lax-Wendroff-Lösung (rechts)

Abbildung 4.4 dargestellte Lösung. Vor der Unstetigkeit ist die tatsächliche Lösung durch Oszillationen, die durch das Differenzenverfahren höherer Ordnung entstehen, verfälscht. Implementieren wir nach jedem Zeitschritt des Lax-Wendroff-Verfahrens eine Filteriteration des Filters (4.16), ergibt sich bei gleicher CFL-Zahl ebenfalls nach  $t = 2.5$  die im rechten Teil von Abbildung 4.4 dargestellte Lösung.

Die Oszillationen sind bis auf sehr kleine Restoszillationen direkt vor der Unstetigkeit geglättet, die Höhe des Stoßes ist erhalten und auch die Stoßlage stimmt mit der tatsächlichen überein, wie eine Überlagerung der gefilterten und ungefilterten Lösung in Abbildung 4.5 links zeigt. Die rechte Darstellung in Abbildung 4.5 zeigt die Lösung des gefilterten Differenzenverfahrens berechnet auf 3200 äquidistanten Gitterpunkten, wodurch die Konvergenz des zusammengesetzten Verfahrens für  $\Delta x \rightarrow 0$  verdeutlicht wird. Erhöhen wir die Anzahl der

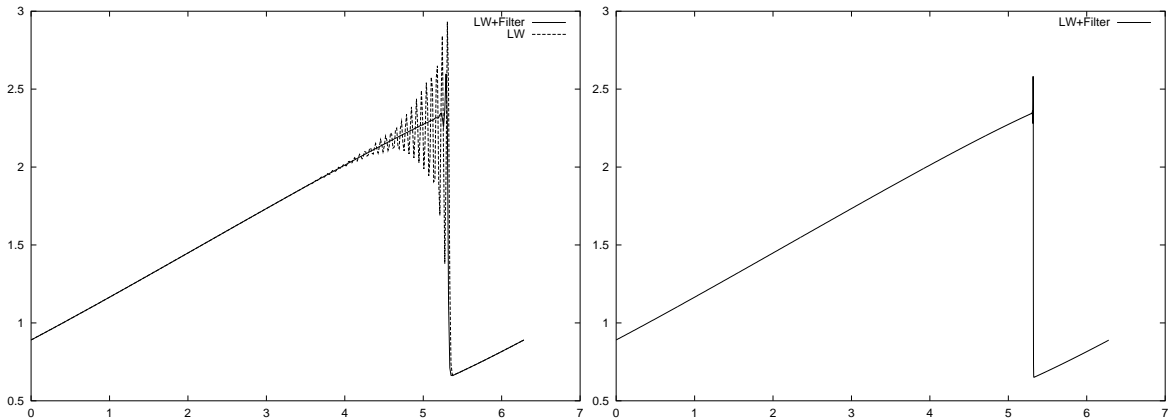


Abbildung 4.5: Vergleich beider Lösungen für  $N = 400$  (links) und gefilterte Lösung für  $N = 3200$  (rechts)

Filteriterationen pro Zeitschritt, verbessert sich die Qualität der Lösung noch weiter. Schon bei drei Filteriterationen nach jedem Schritt des Differenzenverfahrens sind die Oszillationen des Lax-Wendroff-Schemas vollständig geglättet, wie Abbildung 4.6 zeigt, wobei allerdings

zu beachten ist, daß das resultierende Verfahren mit steigender Anzahl von Filteriterationen rechentechnisch etwas an Geschwindigkeit verliert. Bei weiterer Erhöhung der Iterationszahl bleibt die Lösung konstant.

Für die Anwendung auf das Cauchy-Problem (1.49) liefert unsere neue Filterroutine damit sehr gute Ergebnisse.

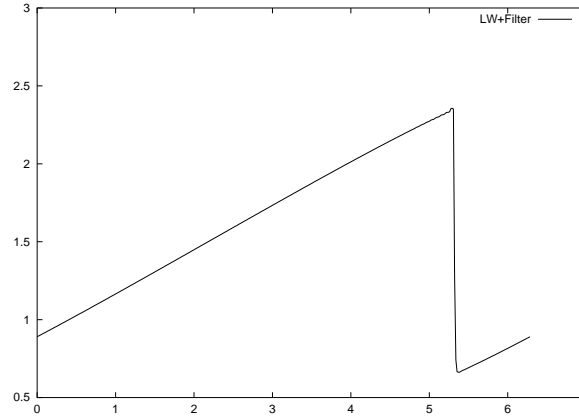


Abbildung 4.6: Lax-Wendroff-Verfahren mit drei Filteriterationen pro Zeitschritt

Nun gehen wir einen Schritt weiter und versuchen, das instabile Zentrale-Differenzen-Verfahren mit Hilfe des diskreten Filters (4.16) zu stabilisieren.

Wir ersetzen das stabile Lax-Wendroff-Verfahren durch das Zentrale-Differenzen-Verfahren

$$U_i^{n+1} = U_i^n - \frac{\Delta t}{2\Delta x}(U_{i+1}^n - U_{i-1}^n)$$

zweiter Ordnung, das bei einer CFL-Zahl von 0.009 auf einem äquidistanten Rechengitter, bestehend aus 400 Gitterpunkten, nach  $t = 2.5$  die in Abbildung 3.9 dargestellte oszillierende Lösung erzeugt. Wenige Zeitschritte später bricht die Lösung aufgrund des instabilen Verfahrens zusammen.

Implementieren wir nach jedem Zeitschritt des Differenzenverfahrens eine Iteration des nicht-linearen Filters (4.16), so ergibt sich dagegen für eine weitaus größere CFL-Zahl von 0.1 die in Abbildung 4.7 links gezeigte Lösung, die zur Veranschaulichung der Filterauswirkungen zusammen mit ihrer ungefilterten Lösung dargestellt ist. Die Oszillationen der zentralen Differenz sind bis auf sehr kleine Schwankungen geglättet, nur kurz vor der Unstetigkeit sind noch etwas stärkere Oszillationen zu beobachten, die sich aber für mehrere Filteriterationen je Zeitschritt vollständig glätten lassen. Die kleinen Schwankungen im linearen Anteil der gefilterten Lösung sind durch das instabile Zentrale-Differenzen-Verfahren verursacht. Ein Vergleich mit der zugehörigen Lax-Wendroff-Lösung zeigt, daß die richtige Lage und Höhe des Stoßes erhalten ist.

Betrachten wir das aus zentraler Differenz und Filterschritt zusammengesetzte Verfahren für kleiner werdene Gitterweite  $\Delta x$ , so stellen wir allerdings fest, daß das Verfahren nicht konvergiert, denn schon bei  $N = 800$ , also halb so großer Gitterweite, entstehen neue Oszillationen in der Lösung, wie Abbildung 4.7 rechts zeigt.

Da diese Probleme bei der Filterimplementierung in das Lax-Wendroff-Verfahren nicht auftreten, weisen die neuen Oszillationen bei  $N = 800$  Gitterpunkten nicht unbedingt auf eine

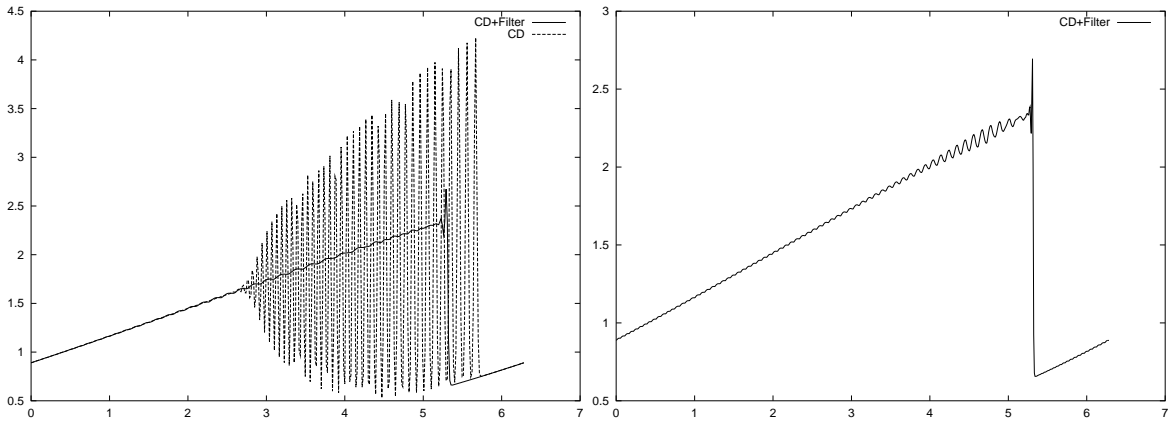


Abbildung 4.7: Gefiltertes und ungefiltertes Zentrale-Differenzen-Verfahren (links), gefilterte Zentrale-Differenzen-Lösung für  $N = 800$  (rechts)

schlecht diskretisierte Filtergleichung hin, sondern sie können durch das instabile Differenzenverfahren verursacht sein. Um genauere Aussagen treffen zu können, untersuchen wir im Folgenden die möglichen Fehlerquellen.

Zunächst besteht die einfache Möglichkeit, daß eine zu groß gewählte CFL-Zahl Ursache für die Divergenz des zusammengesetzten Verfahrens ist. Numerische Experimente zeigen, daß zwar bei  $CFL = 0.01$  noch neue Oszillationen für kleiner werdene Gitterweite auftreten, aber bei  $CFL = 0.009$  tritt dieses Phänomen nicht mehr auf und die Resultate lassen auf die Konvergenz des Verfahrens schließen. Abbildung 4.8 zeigt das Ergebnis für  $N = 400$  und  $N = 3200$  Gitterpunkte.

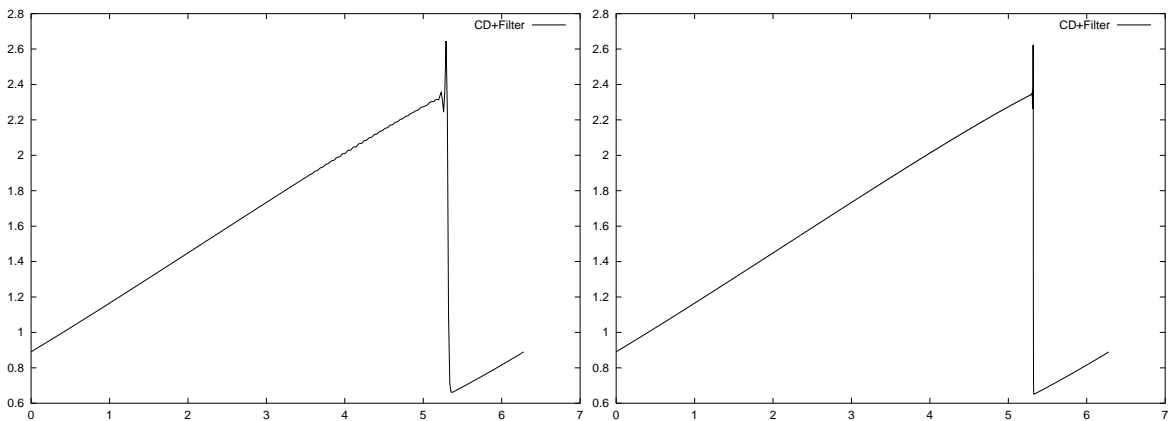


Abbildung 4.8: Lösung des gefilterten Zentrale-Differenzen-Verfahrens bei  $CFL = 0.009$  für  $N = 400$  (links) und  $N = 3200$  (rechts)

Eine weitere Fehlerquelle kann die Anzahl der Filteriterationen pro Zeitschritt sein. Aufgrund des stark oszillierenden Differenzenverfahrens ist es möglich, daß eine Filteriteration zu wenig ist, um das Verfahren für  $\Delta x \rightarrow 0$  zu stabilisieren. Implementieren wir zehn Filteriterationen nach jedem Zeitschritt, ergeben sich für  $N = 400$  und  $N = 6400$  bei einer CFL-Zahl von

0.1 die in Abbildung 4.9 dargestellten Ergebnisse. Die Oszillationen direkt vor der Unstetigkeit sind vollständig geglättet und für  $\Delta x \rightarrow 0$  konvergiert das Verfahren. Genauso gute Ergebnisse ergeben sich bereits bei drei Filteriterationen pro Zeitschritt.

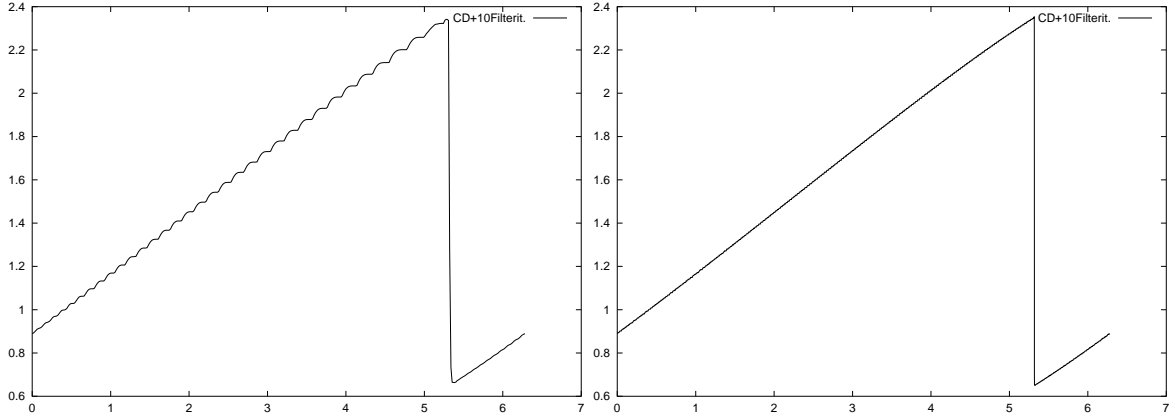


Abbildung 4.9: Gefiltertes Zentrale-Differenzen-Verfahren mit 10 Filterschritten je Zeitschritt für  $N = 400$  (links) und  $N = 6400$  (rechts)

Eine weitere Ursache für zunehmende Oszillationen bei  $\Delta x \rightarrow 0$  kann die Größe der Zeitschrittweite sein. Für hyperbolische Erhaltungsgleichungen wird die Zeitschrittweite  $\Delta t$  mit Hilfe der CFL-Bedingung gemäß  $\Delta t = \text{CFL} \cdot \Delta x / u_{\max}$  bestimmt. Da wir dieses  $\Delta t$  auch bei der Implementierung der Filterroutine verwenden, der Filterschritt allerdings nicht mehr hyperbolischen sondern parabolischen Charakter hat, können sich aufgrund der zu großen hyperbolischen Zeitschrittweite Fehler ergeben. Abhilfe leistet eine Zeitschrittweite, die für das gesamte Verfahren aus der parabolischen Bedingung  $\Delta t = \text{CFL} \cdot \Delta x^2 / u_{\max}$  gewonnen wird. Mit dieser parabolischen Zeitschrittweite ergibt das Zentrale-Differenzen-Verfahren mit einer Filteriteration pro Zeitschritt bei einer CFL-Zahl von 0.5 für  $N = 400$  und  $N = 1600$  die in Abbildung 4.10 dargestellten Lösungen. Bis auf eine Restoszillation direkt vor der Unstetigkeit ist die Lösung sehr gut approximiert, und das Verfahren konvergiert für kleiner werdende Gitterweite  $\Delta x \rightarrow 0$ .

Darüber hinaus besteht die Möglichkeit, eine hyperbolische Zeitschrittweite für das Differenzenverfahren mit einer parabolischen Zeitschrittweite für die Filterroutine zu kombinieren, um das resultierende Verfahren zu beschleunigen. Da die Filterterme durch die parabolische Zeitschrittweite allerdings betragsmäßig sehr klein werden, muß die Anzahl der Filteriterationen gleichzeitig erhöht werden, damit ein stabiles Verfahren entsteht, so daß diese Kombination der Zeitschrittweiten nicht sehr effektiv ist.

Insgesamt zeigt sich, daß das Auftreten neuer Oszillationen bei kleiner werdender Gitterweite verhindert werden kann, indem entweder die CFL-Zahl verkleinert, die Anzahl der Filteriterationen vergrößert oder eine parabolische statt einer hyperbolischen Zeitschrittweite eingeführt wird. Mit diesen zusätzlichen Änderungen ergibt das gefilterte Zentrale-Differenzen-Verfahren eine stabile und konvergente Methode zur Approximation von Lösungen hyperbolischer Erhaltungsgleichungen. Die Implementierung der nichtlinearen Filterroutine (4.16) in die zentrale Differenz stabilisiert das instabile Verfahren, glättet die Oszillationen und erhält die Höhe und Lage der Unstetigkeit. Allerdings ist zu beachten, daß jede angeführte, für die Konvergenz notwendige, Änderung das resultierende Differenzenverfahren gegenüber

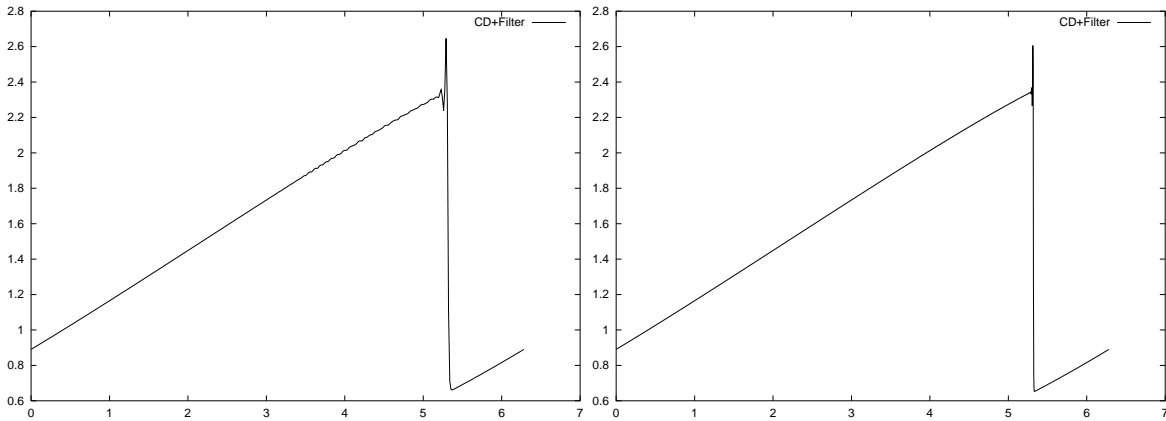


Abbildung 4.10: Gefiltertes Zentrale-Differenzen-Verfahren mit parabolischer Zeitschrittweite für  $N = 400$  (links) und  $N = 1600$  (rechts)

dem gefilterten, divergenten Verfahren verlangsamt. Dabei ist die Erhöhung der Anzahl der Filteriterationen pro Zeitschritt die rechenzeitgünstigste Maßnahme, die Verwendung einer parabolischen Zeitschrittweite kostet dagegen am meisten Rechenzeit. Da das zugrundegelegte Differenzenverfahren, die zentrale Differenz, sehr geringen Rechenaufwand erfordert, ist das gesamte Lösungsverfahren bezüglich der Rechenzeit dennoch relativ effektiv.

### Filterimplementierung für zweidimensionale skalare Probleme

Wir betrachten das zweidimensionale skalare Anfangswertproblem (1.61), das in Kapitel 1 zusammen mit seiner analytischen Lösung beschrieben ist. Eine numerische Lösungsapproximation, berechnet mit dem Lax-Wendroff-Verfahren auf einer Triangulierung aus 5802 Dreiecken, ist in Abbildung 4.11 links als Draufsicht mit Hilfe von Isolinien dargestellt. Entgegen dem Maximum 1.5 und dem Minimum  $-1.0$  der tatsächlichen Lösung ist das Maximum beziehungsweise Minimum der numerischen Lösung 2.0 bzw.  $-1.47$ , und es lassen sich Oszillationen in der Umgebung des Stoßes erkennen.

Bei der Implementierung des diskreten Filters (4.16) auf einem zweidimensionalen Dreiecksgitter ist zu beachten, daß die Filterterme aus Konsistenzgründen nicht mehr, wie in (4.16), mit  $\Delta x$  sondern mit  $\Delta x^2$  multipliziert werden müssen. Außerdem besitzt ein Gitterpunkt nun nicht mehr nur zwei, sondern bis zu sieben Nachbarpunkte, so daß der Algorithmus auf diese Situation angepaßt werden muß. Ob der betrachtete Punkt ein Extremwert ist und der Filter damit zu aktivieren ist, überprüfen wir, indem wir den maximalen und den minimalen Punktwert unter allen Nachbarpunkten berechnen. Ist der Wert des betrachteten Punktes größer als das Maximum oder kleiner als das Minimum, handelt es sich um einen Extremwert und der Filter wird für den Punkt an der Stelle  $i$  über  $d_i = 1$  aktiviert. In einem ersten Durchlauf über alle Punkte des Gitters wird dieser Diffusionsfaktor  $d_i$  für jeden Punkt bestimmt. Der eigentliche Filterschritt wird in einem zweiten Punktedurchlauf ausgeführt, wobei jetzt die Differenzen zu allen Nachbarn berücksichtigt werden. Für einen Gitterpunkt

mit  $l$  Punktnachbarn ist der nichtlineare diskrete Filter durch

$$U_i^{n+1} = U_i^n + \frac{\Delta t}{2} [(d_1 + d_i)(U_1^n - U_i^n) + (d_2 + d_i)(U_2^n - U_i^n) + \dots + (d_l + d_i)(U_l^n - U_i^n)] \quad (4.17)$$

gegeben, wobei  $d_1, \dots, d_l$  die Diffusionsfaktoren der Nachbarnpunkte  $1, \dots, l$  sind und den Wert eins oder Null besitzen, je nachdem, ob der jeweilige Nachbarnpunkt ein Extremwert ist oder nicht. Das heißt, es gilt

$$d_m = \begin{cases} 1 & , \text{ falls Nachbarnpunkt } m \text{ ein Extremwert ist} \\ 0 & , \text{ sonst} \end{cases}$$

für  $m = 1, \dots, l$ . Die Größen  $U_1, \dots, U_l$  bezeichnen die Punktwerte der Punktnachbarn  $1, \dots, l$ .

Damit ist der nichtlineare Filter für die Anwendung auf zweidimensionale Triangulierungen verallgemeinert und wird zunächst auf die oben beschriebene und in Abbildung 4.11 links dargestellte stationäre, oszillierende Lax–Wendroff–Lösung angewendet. Nach 500 Filteriterationen, implementiert nach Beendigung des Differenzenverfahrens, ergibt sich die im rechten Teil von Abbildung 4.11 gezeigte Lösung. Die Oszillationen sind deutlich geglättet, der Stoß ist erhalten und maximaler beziehungsweise minimaler Punktwert liegen bei 1.502 beziehungsweise  $-1.00$ . Im Gegensatz zu der Anwendung des TV–Filters von Chan, Osher und Shen (vergleiche Abbildung 3.2) ist der lineare Teil der Lösung exakt erhalten. Die Filterroutine (4.16) liefert damit sehr gute Resultate beim nichtlinearen Glätten oszillierender Lösungen mit Unstetigkeiten.

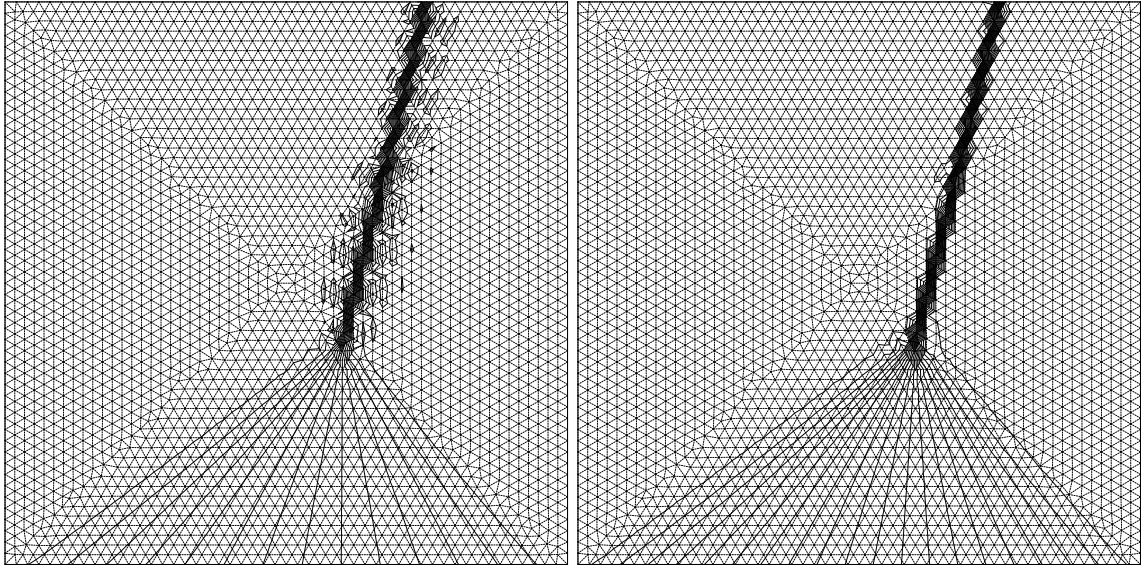


Abbildung 4.11: Oszillierende Lax–Wendroff–Lösung (links) und mit (4.16) gefilterte Lax–Wendroff–Lösung

Wie auch im eindimensionalen Fall ist es von Interesse, den diskreten Filter nach jedem Zeitschritt des Differenzenverfahrens zu implementieren.



Integrieren wir nach jedem Zeitschritt des Lax–Wendroff–Verfahrens 15 Iterationen der Filterroutine (4.17), so ergibt sich die in Abbildung 4.12 dargestellte Lösung. Auch hier sind die

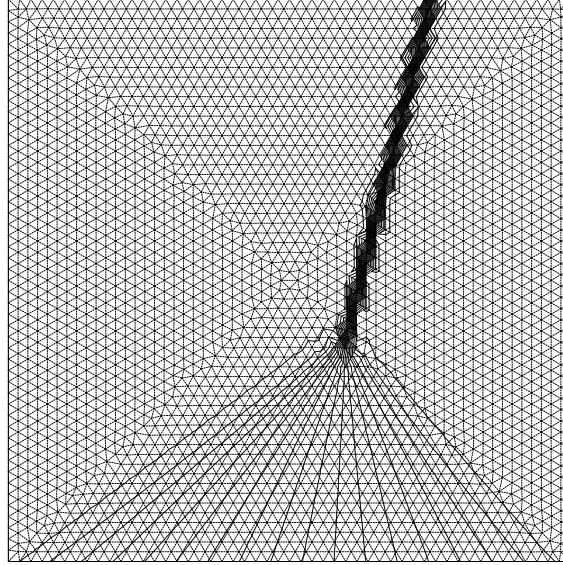


Abbildung 4.12: Lax–Wendroff–Lösung mit 15 Filteriterationen pro Zeitschritt

Oszillationen geglättet und die Lage und Höhe des Stoßes ist erhalten. Maximaler und minimaler Punktwert liegen bei 1.67 beziehungsweise  $-1.14$ , womit das Resultat nicht ganz so gut ist, wie das Ergebnis bei der Implementierung von 500 Filteriterationen nach Beendigung des Differenzenverfahrens. Die Lösung läßt sich durch eine größerer Anzahl von Filteriterationen pro Zeitschritt weiter verbessern, allerdings erhöht sich damit auch der erforderliche Rechenaufwand des resultierenden Verfahrens.

Wie numerische Experimente zeigen, läßt sich die Lösungsqualität auch verbessern, indem die Filterterme in (4.17) mit einem Faktor größer eins multipliziert werden. Da die Filterterme bei der Implementierung in zweidimensionale Probleme aus Konsistenzgründen mit dem Faktor  $\Delta x^2$  multipliziert wurden, liefert der Filter weitaus kleinere Korrekturterme als die ursprüngliche Filterroutine, so daß eine steigende Anzahl von Iterationen pro Zeitschritt nötig ist, um gleiche Korrekturbeträge zu erhalten. Die Multiplikation der Filterterme mit einer Konstanten größer eins wirkt dieser Eigenschaft entgegen und ermöglicht betragsmäßig größere Korrekturen pro Filteriteration. Die Anzahl der Iterationen je Zeitschritt kann damit reduziert werden, das Verfahren wird schneller und die Qualität der Lösung steigt. Der Nachteil dieser Prozedur ist allerdings, daß es ad hoc keine Anhaltspunkte für die Wahl der Größe des konstanten Faktors gibt und er durch Ausprobieren gefunden werden muß. Dabei ist zu beachten, daß ein zu großer Faktor die Lösung des resultierenden Verfahrens auch zerstören kann, wenn die Korrekturterme unverhältnismäßig groß werden.

Wählen wir für unsere Anwendung den Faktor 10.0, so ergibt sich bei vier Filteriterationen nach jedem Zeitschritt des Lax–Wendroff–Verfahrens die in Abbildung 4.13 dargestellte Lösung. Der minimale Wert der Lösung ist  $-1.03$ , der maximale 1.56, so daß dieses Verfahren schneller ist und bessere Resultate liefert als das vorangehende.

Statt des stabilen Lax–Wendroff–Verfahrens betrachten wir im Folgenden das instabile Zentrale–Differenzen–Verfahren, das ohne zusätzliche Filterroutinen keine stationäre Lösung des

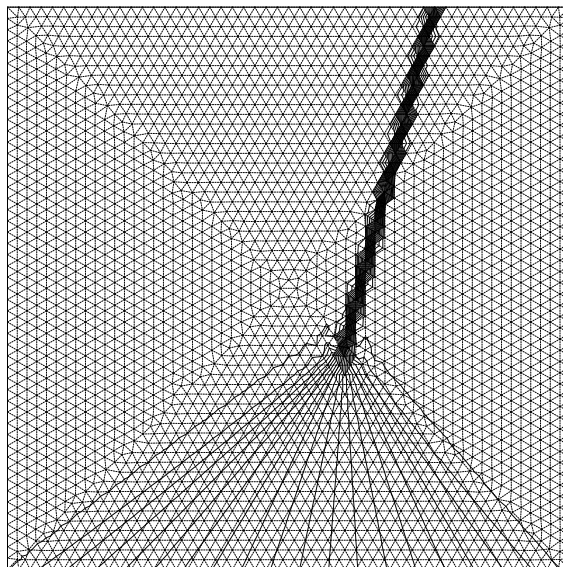


Abbildung 4.13: Lax–Wendroff–Lösung mit vier Filteriterationen pro Zeitschritt und Faktor 10 im Filterterm

vorliegenden Anfangswertproblems liefert, da das Verfahren aufgrund starker Oszillationen schon nach wenigen Zeitschritten zusammenbricht.

Da die Oszillationen der zentralen Differenz weitaus größer sind als die des Lax–Wendroff–Verfahrens, ist zu vermuten, daß jetzt eine größere Anzahl von Filteriterationen pro Zeitschritt nötig ist, um das Verfahren zu stabilisieren. Numerische Experimente bestätigen diese Vermutung. Mit 40 Filteriterationen nach jedem Zeitschritt der zentralen Differenz ergibt sich bei einer CFL–Zahl von 0.9 die in Abbildung 4.14 links dargestellte Lösung. Zwar sind noch Restoszillationen im oberen Teil der Lösung zu erkennen und die eigentlich geradlinigen Isolinien, die den unteren, linearen Teil der Lösung beschreiben, sind fälschlicher Weise stark gekrümmt, aber dennoch hat das Verfahren die tatsächliche Unstetigkeit der Lösung gut approximiert. Maximaler und minimaler Punktwert liegen mit 1.71 und  $-1.08$  nah beim Maximum und Minimum der tatsächlichen Lösung, so daß die noch vorhandenen Oszillationen von kleiner Größenordnung sind. Insgesamt ist das numerische Resultat zwar nicht optimal, aber dennoch beeindruckt die Tatsache, daß das instabile Zentrale–Differenzen–Verfahren, ergänzt durch die nichtlineare Filterroutine (4.17), überhaupt ein stabiles Verfahren darstellt, das die Unstetigkeit der tatsächlichen Lösung gut approximiert.

Auch hier kann versucht werden, den Einfluß des Korrekturbetrags, der durch die Filterterme bestimmt wird, zu vergrößern, indem ein konstanter Faktor größer eins in den Filterterm eingefügt wird. Wählen wir für diese Konstante den Wert 10.0, ergibt sich aus dem zentralen Differenzenverfahren mit 10 Filteriterationen pro Zeitschritt das im rechten Teil von Abbildung 4.14 dargestellte Ergebnis. Die Oszillationen sind geglättet und der Stoß hat sich scharf herausgebildet und stimmt in Höhe und Lage mit der tatsächlichen Unstetigkeit überein. Maximaler und minimaler Punktwert sind 1.55 beziehungsweise  $-1.03$  und stimmen mit dem Maximum und dem Minimum der exakten Lösung fast überein. Nur der linear abfallende Teil der Lösung ist nach wie vor gestört, wie sich an den gekrümmten Isolinien erkennen läßt. Zurückzuführen ist dieser Fehler einerseits auf das Zentrale–Differenzen–Verfahren, das selbst den linearen Anteil der Lösung durch Oszillationen zerstört, andererseits auf die Ei-

genschaft des Filters, daß er nur bei isolierten Extremwerten aktiviert wird und nicht bei in einer Raumrichtung ausgedehnten Extremwerten, die aus mehr als einem Punkt bestehen.

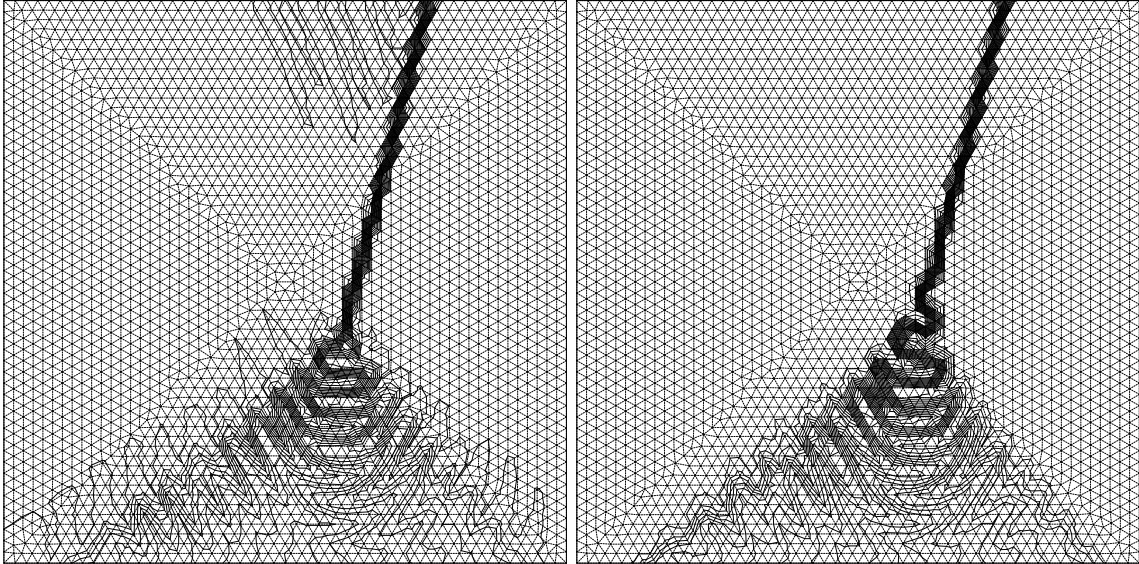


Abbildung 4.14: Zentrale-Differenzen-Verfahren mit 40 Filteriterationen pro Zeitschritt (links) bzw. 10 Filteriterationen mit Faktor 10 im Filterterm

### Filterimplementierung für eindimensionale Systeme von Erhaltungsgleichungen

Im Weiteren wird untersucht, in wie weit die Filterroutine auch für Systeme von Erhaltungsgleichungen eingesetzt werden kann. Wir gehen dazu zunächst zu dem eindimensionalen Fall zurück und betrachten die eindimensionalen Eulergleichungen der Gasdynamik, die bereits in Kapitel 1.3 vorgestellt wurden und in konservativen Variablen durch (1.50) gegeben sind. Für das numerische Lösen dieses hyperbolischen Systems von Erhaltungsgleichungen wurde eine Vielzahl von Finite-Volumen-Methoden entwickelt, von denen wir im Folgenden das Verfahren von Roe näher betrachten wollen. Es dient uns als grundlegendes Lösungsverfahren für die eindimensionalen Eulergleichungen, in das wir die entwickelte Filterroutine integrieren, um ihre Anwendbarkeit für eindimensionale Systeme von Erhaltungsgleichungen zu prüfen.

#### Approximativer Riemann-Löser von Roe

Das Verfahren von Roe gehört zur Klasse der approximativen Riemann-Löser. Wie der Name andeutet, werden die Eulergleichungen dabei nicht mit Hilfe exakter Riemann-Probleme sondern mit einfacheren approximativen Riemann-Problemen gelöst. Statt des exakten Riemann-Problems

$$\frac{\partial U}{\partial t} + \frac{\partial f(U)}{\partial x} = 0, \quad U(0, x) = U_L, \quad x < 0; \quad U(0, x) = U_R, \quad x > 0$$

betrachtet Roe [49] als Approximation des linearen Riemann–Problems

$$\begin{aligned} \frac{\partial U}{\partial t} + A \frac{\partial U}{\partial x} &= 0, \quad A \in \mathbb{R}^{m \times m} \text{ konstant} \\ U(0, x) &= U_L, \quad x < 0; \quad U(0, x) = U_R, \quad x > 0. \end{aligned} \quad (4.18)$$

Die Matrix  $A = A(U_L, U_R)$  hängt dabei von  $U_L$  und  $U_R$  ab und erfüllt die folgenden Bedingungen:

- $f(V) - f(W) = A(V, W)(V - W)$ ,
- $A(V, W) \rightarrow f'(V)$  falls  $W \rightarrow V$ ,
- $A(V, W)$  hat nur reelle Eigenwerte,
- $A(V, W)$  hat ein vollständiges System von Eigenvektoren.

Die erste Bedingung stellt sicher, daß sich die exakte Lösung des tatsächlichen Riemann–Problems ergibt, wenn die Lösung aus einem einzigen Stoß oder einer einzigen Kontaktunstetigkeit besteht, Bedingung zwei sichert die Konsistenz mit der ursprünglichen Gleichung und Bedingung drei und vier garantieren die Lösbarkeit des linearen Riemann–Problems.

Das Verfahren von Roe ist dann durch

$$\begin{aligned} \frac{dU_i}{dt} + \frac{1}{h}(F_{i+1/2} - F_{i-1/2}) &= 0, \\ F_{i+1/2} &= F(U_i, U_{i+1}) = A_{i+1/2}(U_i, U_{i+1})\hat{U}(0; U_i, U_{i+1}) \end{aligned} \quad (4.19)$$

gegeben, wobei  $\hat{U}$  die exakte Lösung des linearen Riemann–Problems (4.18) mit  $U_L = U_i$  und  $U_R = U_{i+1}$  ist und die Matrix  $A$  mit  $A_{i+1/2}$  bezeichnet ist.  $A_{i+1/2}$  heißt Roe–Matrix und  $F_{i+1/2}$ , definiert durch (4.19), wird Roe–Fluß genannt.

Der Roe–Fluß läßt sich weiter umformen zu

$$F_{i+1/2} = \frac{1}{2}(f(U_i) + f(U_{i+1})) - \frac{1}{2} \sum_{p=1}^m |\lambda_p| \alpha_p R_p, \quad (4.20)$$

wobei  $R_1, \dots, R_m$  die  $m$  linear unabhängigen Eigenvektoren von  $A$  zu den korrespondierenden Eigenwerten  $\lambda_1 < \dots < \lambda_m$  sind, und  $\alpha_1, \dots, \alpha_m$  sind implizit definiert durch

$$U_R - U_L = \sum_{p=1}^m \alpha_p R_p.$$

Zur Lösung der Eulergleichungen mit Hilfe des Roe–Verfahrens muß eine Matrix  $A_{i+1/2}$  gefunden werden, die die angeführten Bedingungen von Roe erfüllt. Ihre Konstruktion wird im Folgenden kurz erläutert.

Der Zustands- und Flußvektor der Eulergleichungen (1.50) kann mit Hilfe des Vektors

$$Z = \sqrt{\rho} \begin{pmatrix} 1 \\ u \\ H \end{pmatrix}$$

und den aus der Zustandgleichung für ideale Gase,

$$p = (\gamma - 1)\rho e = (\gamma - 1)(\rho E - \frac{1}{2}\rho u^2),$$

resultierenden Gleichungen

$$p = \frac{\gamma - 1}{\gamma}(\rho H - \frac{1}{2}\rho u^2), \quad \rho E = \frac{1}{\gamma}\rho H + \frac{\gamma - 1}{2\gamma}\rho u^2$$

umgeformt werden zu

$$U = \begin{pmatrix} z_1^2 \\ z_1 z_2 \\ \frac{1}{\gamma} z_1 z_3 + \frac{\gamma - 1}{2\gamma} z_2^2 \end{pmatrix}, \quad f = \begin{pmatrix} z_1 z_2 \\ \frac{\gamma + 1}{2\gamma} z_2^2 + \frac{\gamma - 1}{\gamma} z_1 z_3 \\ z_2 z_3 \end{pmatrix}.$$

Da die Einträge von  $U$  und  $f$  homogene quadratische Funktionen von Elementen von  $Z$  sind, gelten mit den Bezeichnungen  $\delta a = a_{i+1} - a_i$  und  $\bar{a} = \frac{1}{2}(a_i + a_{i+1})$  die Identitäten  $\delta(ab) = \bar{a}\delta b + \bar{b}\delta a$  und  $\bar{a}\delta a = \frac{1}{2}\delta(aa)$ . Mit Hilfe dieser Gleichungen läßt sich die Gültigkeit von

$$\begin{aligned} \delta U &= \bar{B}\delta Z, \quad \bar{B} = \begin{pmatrix} 2\bar{z}_1 & 0 & 0 \\ \bar{z}_2 & \bar{z}_1 & 0 \\ \frac{1}{\gamma}\bar{z}_3 & \frac{\gamma - 1}{\gamma}\bar{z}_2 & \frac{1}{\gamma}\bar{z}_1 \end{pmatrix}, \\ \delta f &= \bar{C}\delta Z, \quad \bar{C} = \begin{pmatrix} \bar{z}_2 & \bar{z}_1 & 0 \\ \frac{\gamma - 1}{\gamma}\bar{z}_3 & \frac{\gamma + 1}{\gamma}\bar{z}_2 & \frac{\gamma - 1}{\gamma}\bar{z}_1 \\ 0 & \bar{z}_3 & \bar{z}_2 \end{pmatrix} \end{aligned}$$

zeigen. Damit gilt

$$\bar{C}\bar{B}^{-1}\delta U = \bar{C}\delta Z = \delta f,$$

so daß

$$A_{i+1/2} = \bar{C}\bar{B}^{-1}$$

die erste Roe-Bedingung erfüllt.

Mit  $R_{i+1/2} := \sqrt{\rho_{i+1}/\rho_i}$  werden die sogenannten Roe-Mittelwerte

$$\bar{u} = \frac{R_{i+1/2}u_{i+1} + u_i}{R_{i+1/2} + 1}, \quad \bar{H} = \frac{R_{i+1/2}H_{i+1} + H_i}{R_{i+1/2} + 1}$$

definiert, wobei  $u_i$  und  $u_{i+1}$  die linken und rechten Geschwindigkeiten bezeichnen. Mit den Roe-Mittelwerten läßt sich die Roe-Matrix umschreiben zu

$$A_{i+1/2} = \begin{pmatrix} 0 & 1 & 0 \\ \frac{\gamma - 3}{2}\bar{u}^2 & (3 - \gamma)\bar{u} & \gamma - 1 \\ \frac{\gamma - 1}{2}\bar{u}^3 - \bar{u}\bar{H} & \bar{H} - (\gamma - 1)\bar{u}^2 & \gamma u \end{pmatrix}. \quad (4.21)$$

Ein Vergleich mit (1.53) zeigt, daß  $A_{i+1/2}$  gerade die Jacobi-Matrix, entwickelt für den Roe-Mittelwert, ist. Daraus folgt sofort, daß  $A_{i+1/2}$  die Roe-Bedingungen zwei bis vier erfüllt. Darüber hinaus zeigen Roe und Pike in [50], daß für die Eulergleichungen die Matrix  $A_{i+1/2}$ , die den Bedingungen eins bis vier genügt, eindeutig ist.

Die Eigenwerte  $\lambda_p$  von  $A_{i+1/2}$ , die für den Roe-Fluß benötigt werden, ergeben sich direkt aus (1.56) und sind

$$\lambda_1 = \bar{u} - \bar{c}, \quad \lambda_2 = \bar{u}, \quad \lambda_3 = \bar{u} + \bar{c},$$

wobei  $\bar{c}$  Schallgeschwindigkeit ist, die sich gemäß

$$\bar{c}^2 = (\gamma - 1)(\bar{H} - \frac{1}{2}\bar{u}^2) \quad (4.22)$$

aus  $\bar{u}$  und  $\bar{H}$  errechnen läßt. Für die Eigenvektoren  $R_p$  ergibt sich aus (1.57) unter Verwendung von (4.22) und einer Skalierung von  $R_1$  und  $R_3$  mit dem Faktor  $-2c/\rho$  bzw.  $2c/\rho$

$$R_1 = \begin{pmatrix} 1 \\ \bar{u} - \bar{c} \\ \bar{H} - \bar{u}\bar{c} \end{pmatrix}, \quad R_2 = \begin{pmatrix} 1 \\ \bar{u} \\ \frac{1}{2}\bar{u}^2 \end{pmatrix}, \quad R_3 = \begin{pmatrix} 1 \\ \bar{u} + \bar{c} \\ \bar{H} + \bar{u}\bar{c} \end{pmatrix}.$$

Die Koeffizienten  $\alpha_p$  im Roe-Fluß ergeben sich aus

$$\sum_{p=1}^m \alpha_p R_p = U_{i+1} - U_i. \quad (4.23)$$

Sei dazu  $R$  die Matrix mit den Spalten  $R_1$ ,  $R_2$  und  $R_3$ . Die Zeilen von  $R^{-1}$ , die wir mit  $R^1$ ,  $R^2$  und  $R^3$  bezeichnen, bilden ein Orthogonalsystem mit den Spalten von  $R$ , das heißt, es gilt

$$R^p \cdot R_q = \delta_q^p,$$

wenn  $\delta_q^p$  das Kroeneckersymbol darstellt. Bilden wir das innere Produkt von (4.23) mit  $R^q$ , erhalten wir

$$\alpha_p = R^q \cdot (U_{i+1} - U_i),$$

wobei

$$\begin{aligned} R^1 &= \left( \frac{\bar{u}}{4\bar{c}}(2 + (\gamma - 1)\frac{\bar{u}}{\bar{c}}), -\frac{1}{2\bar{c}}(1 + (\gamma - 1)\frac{\bar{u}}{\bar{c}}), \frac{\gamma - 1}{2}\frac{1}{\bar{c}^2} \right), \\ R^2 &= \left( 1 - \frac{\gamma - 1}{2}\frac{\bar{u}^2}{\bar{c}^2}, (\gamma - 1)\frac{\bar{u}}{\bar{c}^2}, -(\gamma - 1)\frac{1}{\bar{c}^2} \right), \\ R^3 &= \left( -\frac{\bar{u}}{4\bar{c}}(2 - (\gamma - 1)\frac{\bar{u}}{\bar{c}}), \frac{1}{2\bar{c}}(1 - (\gamma - 1)\frac{\bar{u}}{\bar{c}}), \frac{\gamma - 1}{2}\frac{1}{\bar{c}^2} \right) \end{aligned}$$

gilt. Damit ist die Bestimmung von  $\lambda_p$ ,  $R_p$  und  $\alpha_p$  vollständig beschrieben und der Roe-Fluß kann gemäß (4.20) berechnet werden.

Abschließend gehen wir noch kurz auf eine alternative Formulierung des Roe-Flußes ein.

Sei  $R$  die Matrix mit den Spalten  $R_p$ , dann ist ihre Inverse  $R^{-1}$  die Matrix mit den Zeilen  $R^p$ . Mit dem oben bestimmten  $\alpha$  ergibt sich dann

$$\sum_{p=1}^m |\lambda_p| \alpha_p R_p = R^{-1} |\Lambda| R (U_{i+1} - U_i)$$

mit der Diagonalmatrix  $|\Lambda| = \text{diag}(|\lambda_1|, \dots, |\lambda_m|)$ . Weiterhin gilt  $R^{-1}\Lambda R = A$  und daher setzen wir  $|A| = R^{-1}|\Lambda|R$ . Damit erhalten wir für den Roe-Fluß die kompakte Formulierung

$$F_{i+1/2} = \frac{1}{2}(f(U_i) + f(U_{i+1})) - \frac{1}{2}|A_{i+1/2}|(U_{i+1} - U_i).$$

### Implementierung des Verfahrens von Roe

Wir diskretisieren das Verfahren von Roe (4.19) in der Zeit mit Hilfe der expliziten Euler-Methode und erhalten

$$U_i^{n+1} = U_i^n - \frac{k}{h}(F_{i+1/2}^n - F_{i-1/2}^n),$$

wobei  $h$  die konstante räumliche und  $k$  die zeitliche Schrittweite bezeichnet. Den Roe-Fluß berechnen wir mit Hilfe der Darstellung (4.20). Als Testproblem wählen wir das klassische eindimensionale Riemann-Problem (1.59) auf dem Intervall  $(0, 1)$ , das an der Stelle  $x_0 = 0.3$  eine Unstetigkeit besitzt. Auf einem äquidistanten Gitter aus 100 Gitterpunkten ergibt sich bei einer CFL-Zahl von 0.9 nach  $t = 0.2$  die im linken Teil von Abbildung 4.15 dargestellte Lösung. In der Verdünnungswelle ist ein kleiner Sprung in der Nähe des sonischen Punktes zu erkennen. Er entsteht, wenn einer der Eigenwerte der Jacobi-Matrix Null wird. Dann ist der Roe-Fluß durch

$$F_{i+1/2} = \frac{1}{2}(f(U_i) + f(U_{i+1}))$$

gegeben, was bedeutet, daß lediglich eine zentrale Diskretisierung verwendet wird, so daß es an dieser Stelle wahrscheinlich ist, daß Unstetigkeiten nicht mehr hinreichend geglättet werden.

Abhilfe kann geschaffen werden, wenn das Verfahren von Roe in der Nähe von sonischen Punkten mit zusätzlicher Dissipation ausgestattet wird. Diese Idee stammt von Harten [24], der die Eigenwerte, die sich Null nähern, leicht erhöht, indem er

$$\begin{aligned} |\tilde{\lambda}_p| &= |\lambda_p| \text{ falls } |\lambda_p| \geq \varepsilon; \\ |\tilde{\lambda}_p| &= \frac{1}{2} \left( \frac{\lambda_p^2}{\varepsilon} + \varepsilon \right) \text{ falls } |\lambda_p| < \varepsilon, \quad p = 1, 2, 3 \end{aligned} \quad (4.24)$$

für kleine, positive Werte  $\varepsilon$  setzt. Da die künstliche Vergrößerung der Eigenwerte nur in Kraft tritt, wenn  $|u| \approx c$  gilt, wird (4.24) sonischer Entropie-Fix genannt.

Für  $\varepsilon = 0.5$  ergibt die Wiederholung des vorangehenden Experiments, jetzt mit integriertem Entropie-Fix, die im rechten Teil von Abbildung 4.15 dargestellte Lösung, in der keine Unstetigkeit in der Nähe des sonischen Punktes mehr zu erkennen ist. Die dargestellten Lösungen zeigen, daß der approximative Riemann-Löser von Roe ein Verfahren erster Ordnung ist, das die tatsächliche Lösungsstruktur zu stark glättet.

Um eine qualitativ bessere Lösung zu erhalten, bauen wir im Folgenden ein Verfahren zweiter Ordnung, das Lax-Wendroff-Verfahren, sowie die neu entwickelte Filterroutine in den Roe-Löser ein.

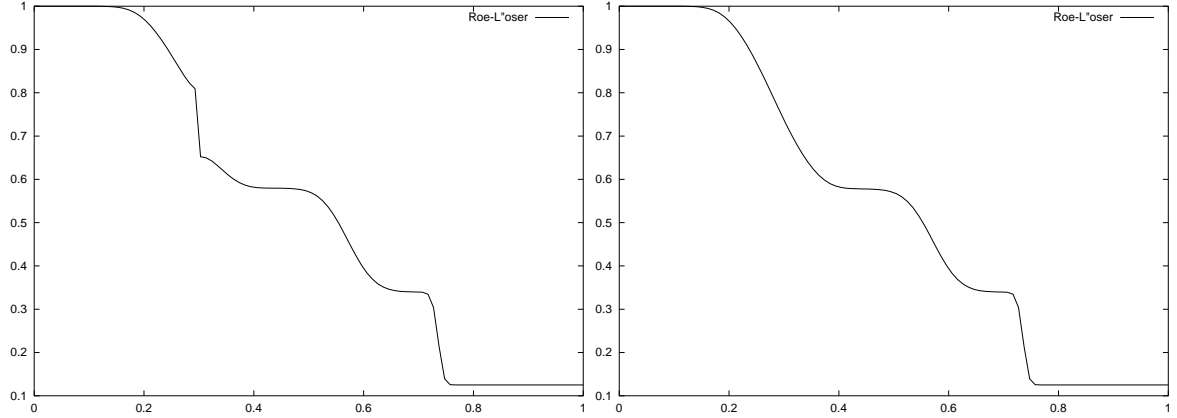


Abbildung 4.15: Lösung des 1D Riemann-Problems (1.59) mit dem Verfahren von Roe ohne (links) und mit (rechts) Entropie-Fix von Harten

### Einbau des Lax-Wendroff-Verfahrens mit Filterroutine

Für den Einbau des Lax-Wendroff-Verfahrens in den Roe-Löser ist lediglich der Roe-Fluß (4.20) durch den Lax-Wendroff-Fluß

$$F_{i+1/2}^{LW} = \frac{1}{2}(f(U_i) + f(U_{i+1})) - \frac{k}{2h} A_{i+1/2}(f(U_{i+1}) - f(U_i))$$

zu ersetzen, wobei  $A_{i+1/2}$  gerade die Roe-Matrix (4.21) ist. Die zeitliche Diskretisierung wird wieder mit dem expliziten Euler-Verfahren ausgeführt.

Um vorerst Probleme an sonischen Punkten auszuschließen, betrachten wir zunächst ein einfacheres Riemann-Problem ohne sonische Punkte, das durch (1.58) gegeben ist und eine Unstetigkeit an der Stelle  $x_0 = 0.3$  besitzt. Auf einem Gitter aus 300 äquidistanten Gitterpunkten ergibt sich bei einer CFL-Zahl von 0.9 nach  $t = 0.2$  die in Abbildung 4.16 links dargestellte Lösung, in der deutliche Oszillationen zu erkennen sind.

Implementieren wir nach jedem Zeitschritt des Differenzenverfahrens den eindimensionalen Filter, indem wir jede Komponente des Gleichungssystems einzeln und unabhängig voneinander filtern, ergibt sich mit nur einer Filteriteration pro Zeitschritt das im rechten Teil von Abbildung 4.16 gezeigte Resultat. Ein großer Anteil der Oszillationen ist geglättet, und die Unstetigkeiten sind in Lage und Höhe erhalten. Mit zunehmender Anzahl von Filteriterationen pro Zeitschritt läßt sich das resultierende Verfahren weiter verbessern, wie in Abbildung 4.17 für fünf und zehn Filteriterationen zu erkennen ist.

Als nächstes betrachten wir das Riemann-Problem (1.59), bei dessen Lösung ein sonischer Punkt entsteht, der in der einfachen Lax-Wendroff-Lösung in Abbildung 4.18 links klar zu erkennen ist. Da bei der Implementierung des Lax-Wendroff-Verfahrens über die Matrix  $A_{i+1/2}$  die Eigenwerte nicht unmittelbar erkennbar sind, kann der Entropie-Fix von Harten nicht verwendet werden. Allerdings berechnet sich  $A_{i+1/2}$  aus

$$A_{i+1/2} = R^{-1} \Lambda R,$$



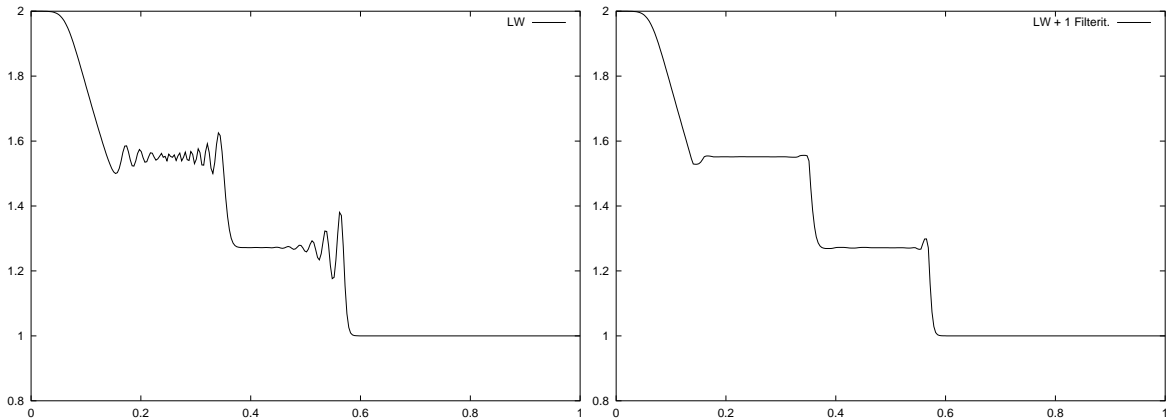


Abbildung 4.16: Ungefilterte (links) und mit einer Iteration pro Zeitschritt gefilterte Lax-Wendroff-Lösung

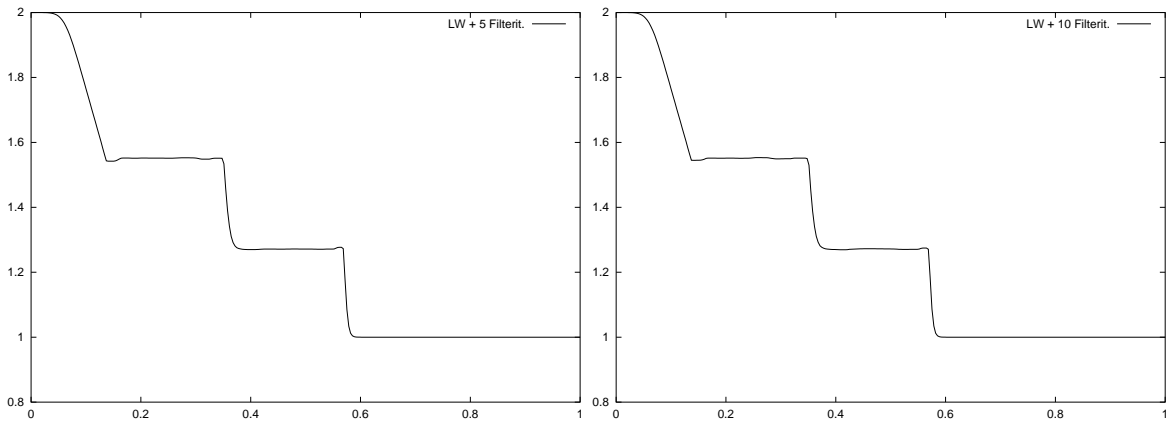


Abbildung 4.17: Gefilterte Lax-Wendroff-Lösung mit 5 (links) bzw. 10 Filteriterationen

wobei  $\Lambda$  die Diagonalmatrix der Eigenwerte ist, so daß bei dieser Formulierung von  $A_{i+1/2}$  die Eigenwerte an dieser Stelle sichtbar sind und der Entropie-Fix dort eingesetzt werden kann. Numerische Experimente zeigen aber, daß der so eingebaute Entropie-Fix von Harten nicht stark genug ist, um die Störungen am sonischen Punkt abzufangen. Da die Eigenwerte nicht nur in der Diagonalmatrix auftreten, sondern auch in  $R$  und  $R^{-1}$ , wo sie bei konservativer Formulierung nicht direkt erkennbar sind, berechnen wir die Matrix  $A_{i+1/2}$  in einem weiteren Versuch mit Hilfe der primitiven Variablen. Der Vorteil bei der Berechnung in primitiven Variablen ist, daß die Eigenwerte stets sichtbar sind und sie daher überall mit Hilfe des Entropie-Fixes abgefangen werden können, wenn sie sich der Null nähern. Allerdings liefert der Entropie-Fix selbst bei dieser Implementierung des Lax-Wendroff-Verfahrens nicht genügend Diffusion, um die Unstetigkeit am sonischen Punkt zu verhindern, wie Abbildung 4.18 rechts zeigt.

Integrieren wir in das Lax-Wendroff-Verfahren mit Entropie-Fix zwei Filteriterationen der nichtlinearen Filterroutine, ergibt sich das in Abbildung 4.19 links dargestellte Resultat. Das Ergebnis läßt sich für eine größere Anzahl von Filteriterationen weiter verbessern, wie der rechte Teil der Abbildung für 10 Iterationen zeigt. Bemerkenswert ist, daß nicht nur die Os-

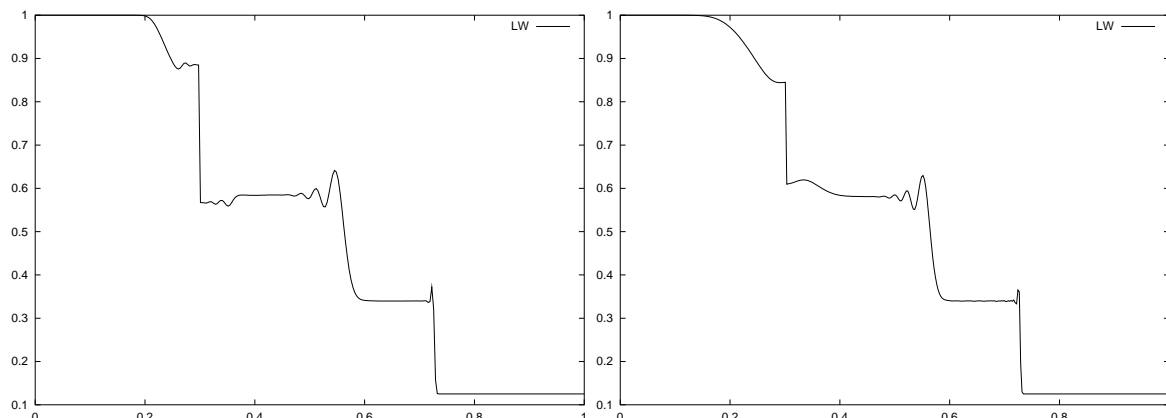


Abbildung 4.18: Lax-Wendroff-Lösung für das Riemann-Problem (1.59) ohne (links) und mit Entropie-Fix

zillationen des Lax-Wendroff-Verfahrens weitgehend geglättet sind, sondern der Filter auch die Störungen am sonischen Punkt stark verringert, was der Entropie-Fix von Harten nicht leistet.

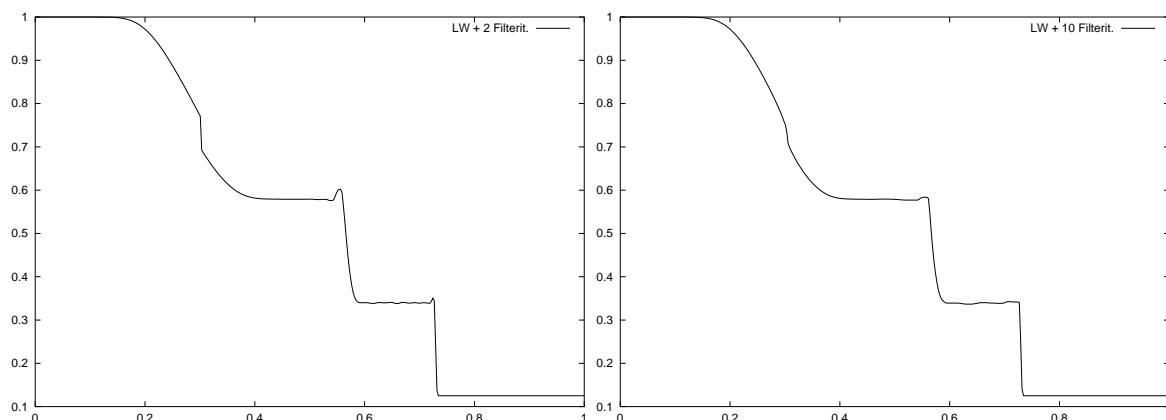


Abbildung 4.19: Mit zwei (links) bzw. 10 Iterationen gefilterte Lax-Wendroff-Lösung für das Riemann-Problem (1.59) mit Entropie-Fix

Noch eindrucksvoller zeigt sich der positive Einfluß des Filters auf Störungen an sonischen Punkten am Testfall von Sod. Er ist gegeben durch die linken und rechten Zustände (1.60) mit einer Unstetigkeit an der Stelle  $x_0 = 0.5$ . Simulieren wir dieses Testproblem mit dem Lax-Wendroff-Verfahren ohne weiteren Entropie-Fix, entstehen schon nach wenigen Zeitschritten sehr starke Oszillationen am unteren Teil der Unstetigkeit, wie Abbildung 4.20 links für  $t = 0.008$  zeigt, und wenig später bricht das Verfahren vollständig zusammen. Implementieren wir zusätzlich zwei Iterationen der nichtlinearen Filterroutine nach jedem Zeitschritt des Lax-Wendroff-Verfahrens, erhalten wir dagegen nach  $t = 0.25$  die im rechten Teil von Abbildung 4.20 dargestellte Lösung. Das Verfahren ist durch den Filtereinbau stabilisiert und Unstetigkeiten aufgrund des sonischen Punktes sind nicht zu erkennen. Zusätzlich sind die Oszillationen des Lax-Wendroff-Verfahrens fast vollständig geglättet. Wie ein Vergleich mit der

exakten Lösung des Sod-Problems in Abbildung 4.21 zeigt, ist die Qualität der numerischen Lösung sehr gut, und Höhe und Lage der Unstetigkeiten sind erhalten.

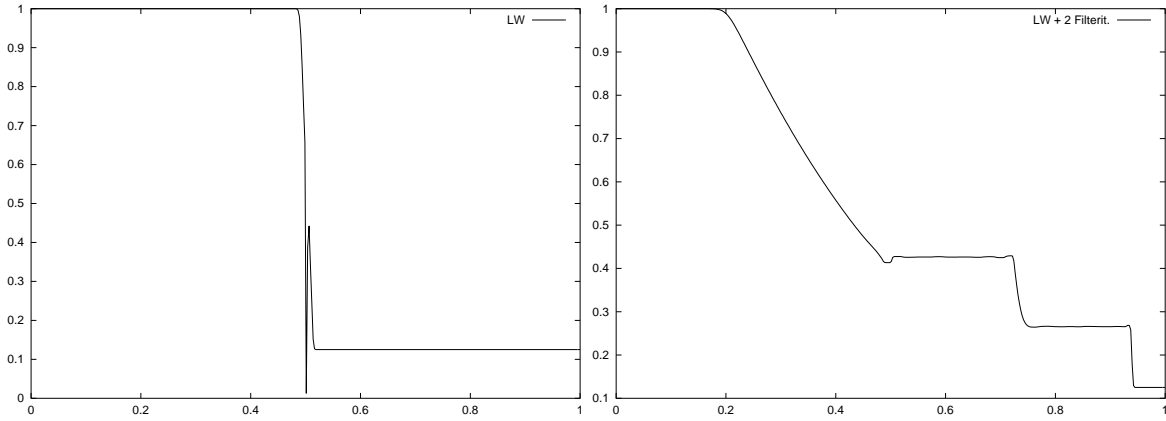


Abbildung 4.20: Instabile Lax-Wendroff-Lösung für das Sod-Problem nach  $t = 0.008$  (links) und Lax-Wendroff-Lösung mit integriertem Filter nach  $t = 0.25$  (rechts)

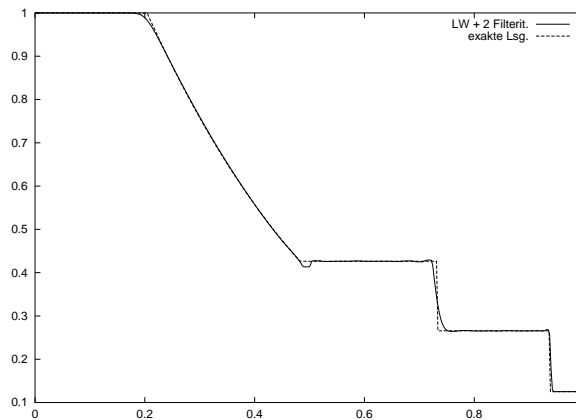


Abbildung 4.21: Vergleich von exakter Lösung und Lax-Wendroff-Lösung mit integriertem Filter

Die entwickelte Filterroutine kann damit nicht nur zum Glätten der Oszillationen von Differenzenverfahren höherer Ordnung eingesetzt werden, sondern dient ebenfalls als geeigneter und einfacher Entropie-Fix bei der Lösung von Riemann-Problemen mit sonischen Punkten. Da der nichtlineare Filter parameterfrei ist, ist seine Anwendung weitaus einfacher als viele in der Literatur beschriebene Methoden, die mit zusätzlicher künstlicher Diffusion Unstetigkeiten in der Umgebung von sonischen Punkten abfangen. Auch die gängige Kombination verschiedener Differenzenverfahren zum Beheben von Problemen an sonischen Punkten ist meistens aufwendiger als die Filterimplementierung, da geeignete Funktionen gefunden werden müssen, die von einem Verfahren zum anderen umschalten.

### Experimentelle Konvergenzordnung

Um Aussagen über die Konvergenzordnung des aus Filter- und Lax-Wendroff-Schritt zu-

sammengesetzten Verfahrens treffen zu können, bestimmen wir im Folgenden die experimentelle Konvergenzordnung (EOC) des neuen Differenzenverfahrens bei Anwendung auf das Riemann–Problem (1.58). Da die exakte Lösung dieses Riemann–Problems bekannt ist, läßt sich die Abweichung von numerischer und exakter Lösung und damit die experimentelle Konvergenzordnung berechnen [32].

Sei  $u_h$  die numerische Lösung, die sich aus der exakten Lösung und einem numerischen Fehler, der von der räumlichen Gitterweite  $h$  abhängt, zusammensetzt. Dann läßt sich  $u_h$  als asymptotische Entwicklung

$$u_h = u + u_1 h^\alpha + \dots$$

darstellen, wobei  $\alpha$  zu bestimmen ist.

Auf einem größeren Gitter der Gitterweite  $2h$  ist  $u_{2h}$  durch die Reihenentwicklung

$$u_{2h} = u + u_1 (2h)^\alpha + \dots$$

gegeben, so daß sich die experimentelle Konvergenzordnung in erster Näherung gemäß

$$\text{EOC} := \alpha = \frac{\ln \left( \frac{\|u - u_h\|_{L^1}}{\|u - u_{2h}\|_{L^1}} \right)}{\ln \left( \frac{1}{2} \right)}$$

berechnen läßt. Die diskrete  $L^1$ –Norm zur Zeit  $t_n$  ist dabei durch

$$\|u - u_h\|_{L^1} = h \sum_i |u(x_i, t_n) - u_h(x_i, t_n)|$$

gegeben.

Tabelle 4.1 stellt die numerischen Fehler und die experimentelle Konvergenzordnung des ungefilterten und des mit zwei Filteriterationen mit (4.16) pro Zeitschritt gefilterten Lax–Wendroff–Verfahrens gegenüber. Da Verfahren höherer Ordnung zumindest für glatte Lösungen typischer Weise eine experimentelle Konvergenzordnung größer eins aufweisen, ist in der dritten Spalte der Tabelle der numerische Fehler und der EOC–Wert für das mit zwei Iterationen gefilterte Lax–Wendroff–Verfahren dargestellt, wobei für die Berechnung die Werte an den beiden Unstetigkeitsstellen jeweils auf einem Intervall der Länge  $\Delta I = 0.023$  weggelassen wurden. Gegenüber dem Lax–Wendroff–Verfahren weist das neue Lösungsverfahren für alle

h	Lax–Wendroff $\ u - u_h\ _{L^1}$ EOC	LW+2 Filterit. $\ u - u_h\ _{L^1}$ EOC	LW+ Filter (ohne Unstetigk.) $\ u - u_h\ _{L^1}$ EOC
0.02	0.032013	0.023692	
0.01	0.021977 0.54	0.012348 0.94	0.008256 0.90
0.005	0.014286 0.62	0.006469 0.93	0.003505 1.24
0.0025	0.009061 0.66	0.003614 0.84	0.001629 1.12
0.00125	0.005299 0.77	0.001879 0.94	0.000839 0.98

Tabelle 4.1: Experimentelle Konvergenzordnung des ungefilterten und gefilterten Lax–Wendroff–Verfahrens

dargestellten Gitterweiten einen wesentlich kleineren numerischen Fehler und eine wesentlich größere experimentelle Konvergenzordnung auf. Werden die Werte in direkter Umgebung der Unstetigkeiten bei der EOC-Berechnung weggelassen, ergeben sich mit Ausnahme der Gitterweiten  $h = 0.01$  und  $h = 0.00125$  EOC-Werte größer eins, so daß das resultierende Verfahren, abgesehen von diesen beiden Gitterweiten, mit anderen hochentwickelten Verfahren höherer Ordnung zu vergleichen ist.

### Filterimplementierung in das Zentrale-Differenzen-Verfahren

Wie bei vorherigen Anwendungen ist es auch hier im beschränkten Umfang möglich, das stabile Lax-Wendroff-Verfahren durch das instabile Zentrale-Differenzen-Verfahren zweiter Ordnung zu ersetzen. Für das Riemann-Problem (1.58) ohne sonische Punkte ergibt die zentrale Differenz, erweitert durch zwei Filteriterationen pro Zeitschritt, bei einer CFL-Zahl von 0.9 die in Abbildung 4.22 links dargestellte Lösung, die sich mit zunehmender Anzahl von Filteriterationen noch weiter verbessern läßt, wie im rechten Teil der gleichen Abbildung für zehn Filteriterationen zu erkennen ist.

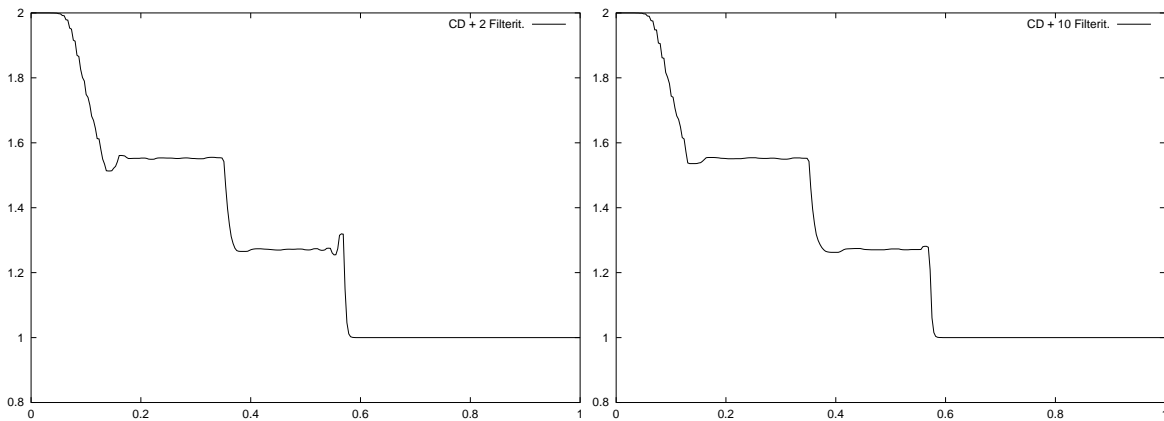


Abbildung 4.22: Lösung des Riemann-Problems ohne sonische Punkte mit dem Zentrale-Differenzen-Verfahren und 2 (links) bzw. 10 (rechts) Filteriterationen pro Zeitschritt

Für die Testfälle von Toro (1.59) und Sod (1.60), bei denen sonische Punkte entstehen, ist der Filter dagegen nicht mehr stark genug, um die zentrale Differenz zu stabilisieren, so daß keine stabilen Verfahren mehr entstehen.

Nachdem die neue Filterroutine erfolgreich auf skalare ein- und zweidimensionale Problemstellungen und auf Systeme von Erhaltungsgleichungen in einer Raumdimension angewendet wurde, untersuchen wir im Folgenden die Filteranwendung auf Systeme von Erhaltungsgleichungen in zwei Raumdimensionen.

### Filterimplementierung für die zweidimensionalen Eulergleichungen

Als Testfall für zweidimensionale Systeme von Erhaltungsgleichungen simulieren wir, wie bereits in Kapitel 3, die Umströmung des NACA0060 Profils. Das Testproblem ist in Kapitel 1.3.2 genau beschrieben, die zweidimensionalen Eulergleichungen sind dort unter (1.62) in

konservativer Formulierung gegeben.

Zunächst betrachten wir als zugrundeliegendes Differenzenverfahren das stabile Lax–Wendroff–Verfahren zweiter Ordnung, das wir im Rahmen einer Finite–Volumen–Approximation auf dem Sekundärgitter der Triangulierung implementieren. Es ergibt sich die in Abbildung 3.6 dargestellte numerische Lösung.

Implementieren wir eine Filteriteration des nichtlinearen Filters (4.17) nach jedem Zeitschritt des Differenzenverfahrens unabhängig für jede Vektorkomponente, so ergibt sich die in Abbildung 4.23 links gezeigte Lösung. Die Oszillationen hinter dem oberen und unterem Stoß sind weitgehend geglättet, ohne daß die Stöße selbst in Höhe oder Lage beeinträchtigt sind. Bei steigender Anzahl von Filteriterationen pro Zeitschritt kann das Resultat weiter verbessert werden, wie die entsprechende Lösung für jeweils 10 Filteriterationen, dargestellt im rechten Teil derselben Abbildung, zeigt. In der Lösung sind keine Oszillationen mehr zu erkennen.

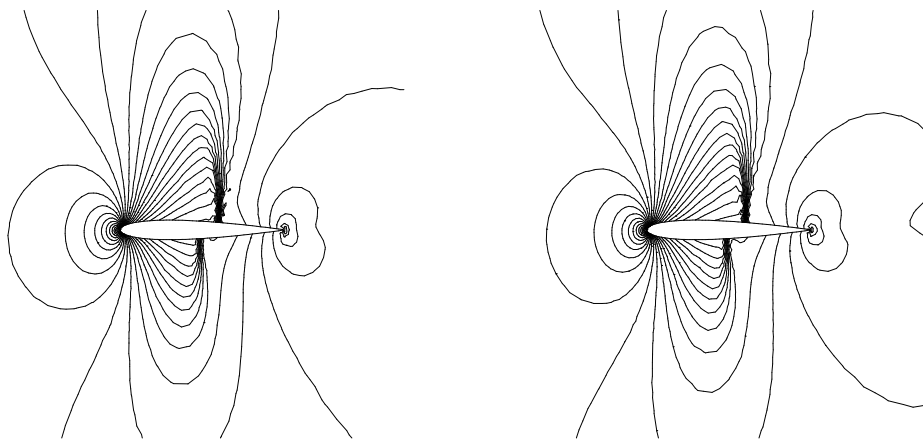


Abbildung 4.23: Lax–Wendroff–Lösung mit einem (links) bzw. 10 (rechts) Filterschritten nach jedem Zeitschritt

Auch hier besteht die Möglichkeit, das resultierende Differenzenverfahren zu beschleunigen, indem wir die Filterterme mit einem Faktor größer eins multiplizieren und nur eine Filteriteration pro Zeitschritt durchführen. Für den Faktor 10.0 ergibt sich beispielweise das in Abbildung 4.24 dargestellte Ergebnis, das weitgehend mit dem Resultat für zehn Filteriterationen ohne zusätzlichen Faktor übereinstimmt.

Ebenso wie in den anderen Testfällen besteht auch hier die Möglichkeit, das stabile Lax–Wendroff–Verfahren durch eine zentrale Differenz zweiter Ordnung zu ersetzen, in der Hoffnung, zusammen mit dem nichtlinearen Filter ein schnelles und rechentechnisch günstiges stabiles Verfahren zu erhalten. Numerische Experimente zeigen, daß das Zentrale–Differenzen–Verfahren schon nach wenigen Zeitschritten so große Oszillationen erzeugt, daß die Filtergleichung (4.17) selbst bei 50 Iterationen nach jedem Zeitschritt nicht genügend Diffusion zur Glättung der Oszillationen liefert, um das resultierende Verfahren zu stabilisieren. Multiplizieren wir dagegen die Filterterme in (4.17) mit einem positiven konstanten Faktor der Größenordnung 10 und integrieren wir diese Filtergleichung nach jedem Zeitschritt in das Zentrale–Differenzen–Verfahren, so ist es möglich, stabile Lösungen zu erzeugen. Für den Faktor 10.0 ergibt die zentrale Differenz mit 10 Filteriterationen pro Zeitschritt beispielsweise das in Abbildung 4.25 links dargestellte Resultat. Die Isolinien weisen zwar in sich eine

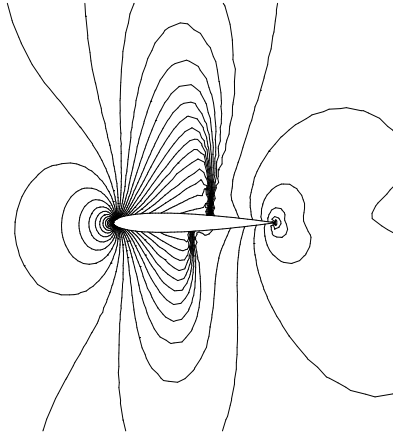


Abbildung 4.24: Lax-Wendroff-Lösung mit einer Filteriteration nach jedem Zeitschritt und Faktor 10 in den Filtertermen

unphysikalisch gezackte Struktur auf, aber dennoch haben sich die beiden verschieden starken Stöße oberhalb und unterhalb des Profils klar herausgebildet. Da das zugrundegelegte Zentrale-Differenzen-Verfahren zweiter Ordnung instabil ist und eine sinnvolle numerische Lösung der Eulergleichungen mit diesem Differenzenverfahren allein unmöglich ist, ist dies eine beachtliche Leistung der nichtlinearen Filtergleichung.

Das zusammengesetzte Lösungsverfahren läßt sich in Bezug auf seine Rechenlaufzeit weiter optimieren, indem die Anzahl der Filteriterationen bei gleichzeitiger Faktorvergrößerung herab gesetzt wird. So ergeben lediglich zwei Filteriterationen pro Zeitschritt mit einem Faktor der Größe 50.0 die im rechten Teil von Abbildung 4.25 dargestellte Lösung, die sich kaum von der vorangehenden unterscheidet.

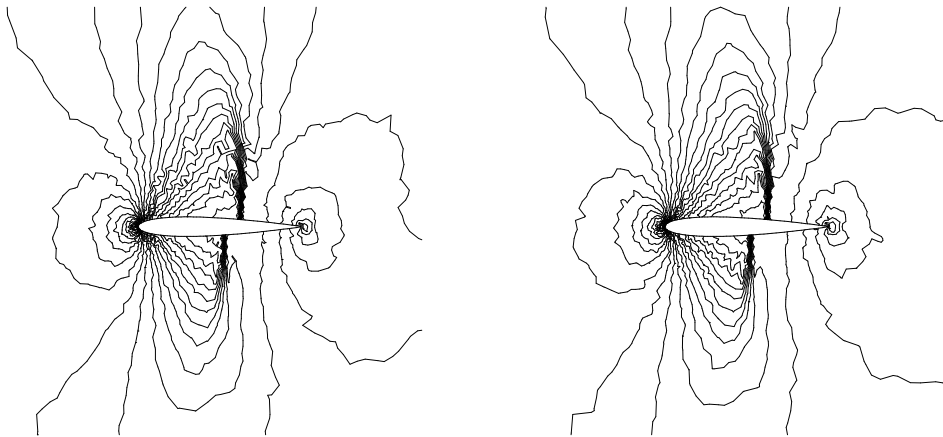


Abbildung 4.25: Zentrale-Differenzen-Lösung mit 10 Filteriterationen und einem Faktor von 10 (links) bzw. mit 2 Iterationen und dem Faktor 50 (rechts)

Aufgrund der gezackten Isolinien, die bei der Filteranwendung auf das Zentrale-Differenzen-Verfahren entstehen, haben wir die regularisierten Versionen

$$u_t = \delta_m(u_x)u_{xx} \quad \text{mit} \quad \delta_m(u_x) = \frac{2}{\pi} \frac{m}{1 + m^2 u_x^2}$$

und

$$u_t = (\phi(u_x)u_x)_x, \text{ mit } \phi(u_x) = \frac{\ln(1 + m^2 u_x^2)}{m\pi|u_x|} + 1 - \frac{2}{\pi} \arctan(m|u_x|), \quad m \in \mathbb{N},$$

der Filtergleichung  $u_t = ((1 - \operatorname{sgn}(|u_x|))u_x)_x$  entwickelt, diskretisiert und nach jedem Zeitschritt in die zentrale Differenz integriert. Die numerischen Ergebnisse zeigen jedoch, daß die regularisierten Filtergleichungen entweder zu viel oder zu wenig Diffusion erzeugen, so daß sich keine qualitativ guten Approximationen ergeben. Darüber hinaus hängt die Filterwirkung in beiden Fällen stark von der Diskretisierung von  $u_x$  ab, und für ungünstige Diskretisierungen ist eine Verschiebung der Stoßlage zu beobachten.

### Zusammenfassung der numerischen Ergebnisse

Insgesamt zeigen die numerischen Experimente, daß die in diesem Unterkapitel neu entwickelte Filtergleichung (4.16) beziehungsweise (4.17) für ein- und zweidimensionale skalare und vektorielle Testprobleme einsetzbar ist, um Oszillationen von Differenzenverfahren höherer Ordnung zu glätten, ohne dabei die physikalischen Stöße zu zerstören. Ein großer Vorteil gegenüber vielen existierenden Methoden, zusätzliche Diffusion in Differenzenverfahren zu integrieren, ist die Parameterlosigkeit unseres diskreten Filters, so daß bei der Anwendung auf verschiedene Problemstellungen eine mühsame Parameterbestimmung, schlimmstenfalls durch Ausprobieren, wegfällt. Darüber hinaus ist es beachtlich, daß der nichtlineare Filter nicht nur Oszillationen stabiler Differenzenverfahren glättet, sondern selbst bei der Anwendung auf instabile Differenzenverfahren zu stabilen und teilweise sehr guten numerischen Lösungen führt.

Die numerischen Ergebnisse für die Stoßrohrprobleme mit sonischen Punkten zeigen, daß der nichtlineare Filter auch als geeigneter Entropie-Fix eingesetzt werden kann und für das Lax-Wendroff-Verfahren zu besseren Ergebnissen als der Entropie-Fix von Harten führt. Für das Stoßrohrproblem von Sod (1.60), für das das Lax-Wendroff-Verfahren schon nach wenigen Zeitschritten instabil wird, führt die Implementierung von zwei Filteriterationen nach jedem Zeitschritt des Lax-Wendroff-Verfahrens zu stabilen Ergebnissen hoher Approximationsgenauigkeit.

Mit Hilfe des nichtlinearen Filters ist es gelungen, sehr einfache und rechentechnisch schnelle Verfahren für das Lösen verschiedenster hyperbolischer Erhaltungsgleichungen zu konstruieren.

#### 4.1.6 Amplituden- und Phasenverhalten der nichtkonservativen Filtergleichung

Die numerischen Experimente zeigen, daß bei der Implementierung der diskreten nichtkonservativen Filtergleichung (4.14), ebenso wie beim TV-Filter von Chan, Osher und Shen, eine Verschiebung der Stoßlage in der numerischen Lösung entsteht. Im Gegensatz zu den bisherigen Filterroutinen, die dieses Fehlverhalten aufweisen, ist der diskrete Filter (4.14) sehr einfach aufgebaut, so daß er mit Hilfe der von Neumannschen Stabilitätsanalyse in Hinblick auf Amplituden- und Phasenverhalten untersucht werden kann. Von den Ergebnissen der



Analyse erhoffen wir uns Aufschluß darüber, ob es, abgesehen von der Nichtkonservativität, weitere Ursachen für das Entstehen des Verschiebungsfehlers gibt.

Die von Neumann Analyse ist eine Fourieranalyse, mit der die Stabilität eines numerischen Finite-Differenzen-Verfahrens für lineare Differentialgleichungen mit konstanten Koeffizienten und periodischen Randbedingungen untersucht werden kann. Dabei wird das Wachstumsverhalten der Rundungsfehler, die durch Rechnerungenauigkeiten entstehen, analysiert. Ergibt die Fourieranalyse, daß diese Rundungsfehler nicht wachsen, ist das numerische Verfahren schwach stabil und strenge Stabilität kann gefolgert werden.

### Grundlagen der von Neumann Analyse

Bevor wir die von Neumann Analyse für den diskreten Filter (4.14), der durch

$$U_j^{n+1} = U_j^n + \frac{\Delta t}{\Delta x} d_j (U_{j+1}^n - 2U_j^n + U_{j-1}^n), \quad (4.25)$$

gegeben ist, durchführen, erläutern wir im Folgenden kurz die allgemeine Vorgehensweise. Dabei orientieren wir uns an den Darstellungen der Stabilitätsanalyse wie sie bei Anderson, Tannehill und Pletcher [3] beziehungsweise Hirsch [26] zu finden sind.

Um Verwechslungen mit der komplexen Zahl  $i$  zu vermeiden, bezeichnen wir die diskreten Gitterpunkte in diesem Abschnitt mit  $(x_j, t_n)$  und die Approximation von  $u(x_j, t_n)$  mit  $U_j^n$ .

Bezeichnet  $\varepsilon$  den Fehler einer numerischen Lösung eines Finite-Differenzen-Verfahrens, der durch rechnerinterne Rundungseffekte entsteht, so läßt sich die numerische Lösung einer Differenzengleichung als

$$N = D + \varepsilon$$

darstellen, wobei  $D$  die exakte Lösung der Differenzengleichung und  $N$  die mit einem realen Computer, also endlicher Genauigkeit, berechnete Lösung ist. Da die numerische Lösung  $N$  die Differenzengleichung erfüllen muß, gilt das gleiche sowohl für die exakte Lösung  $D$  als auch für den Fehler  $\varepsilon$ , so daß  $D$  und  $\varepsilon$  die gleichen zeitlichen Wachstumseigenschaften besitzen. Wir setzen voraus, daß der Fehler  $\varepsilon(x, t)$  als Reihenentwicklung in der Form

$$\varepsilon(x, t) = \sum_m b_m(t) e^{ik_m x} \quad (4.26)$$

dargestellt werden kann, wobei wir als Periode der Grundfrequenz ( $m = 1$ ) die Länge  $2L$  wählen. Bezeichnet  $M$  die Anzahl der Intervalle mit Länge  $\Delta x$ , die in der Länge  $L$  enthalten sind, so ist die Wellenzahl  $k_m$  gegeben durch

$$k_m = \frac{m\pi}{L}, \quad m = 0, 1, 2, \dots, M,$$

und die zugehörigen Frequenzen sind

$$f_m = \frac{k_m}{2\pi} = \frac{m}{2L}.$$

Die Frequenz gibt dabei an, wieviele Wellenlängen in jedem Längenintervall  $2L$  enthalten sind. Für  $m = 0$  gilt  $f_0 = 0$ , was einem konstanten Term in der Reihenentwicklung entspricht.

Da die zugrundeliegende Differenzengleichung linear ist, gilt das Superpositionsprinzip, so daß das Verhalten eines einzelnen Summanden

$$\varepsilon_m(x, t) = b_m(t)e^{ik_mx}$$

der Reihe (4.26) analysiert werden kann. Gesucht sind Lösungen der Form

$$z^n e^{ik_mx},$$

die sich für  $t = 0$  auf den Ausdruck  $e^{ik_mx}$  reduzieren. Dies gewährleistet

$$z = e^{a\Delta t},$$

und mit

$$z^n = e^{an\Delta t} = e^{at}$$

ergibt sich

$$\varepsilon_m(x, t) = e^{at} e^{ik_mx}, \quad (4.27)$$

wobei  $k_m$  reell ist,  $a$  aber komplex sein kann. Da der Fehler  $\varepsilon$  und damit auch  $\varepsilon_m$  die Differenzengleichung erfüllt, können wir die rechte Seite von (4.27) in die Differenzengleichung einsetzen und mit Hilfe der Beziehungen

$$\cos \beta = \frac{e^{i\beta} + e^{-i\beta}}{2}, \quad \sin \beta = \frac{e^{i\beta} - e^{-i\beta}}{2i}$$

nach  $e^{a\Delta t}$  auflösen.

Darüber hinaus gilt für jede Frequenz, die in der Lösung der Differenzengleichung für den Fehler enthalten ist,

$$\varepsilon_j^{n+1} = e^{a\Delta t} \varepsilon_j^n.$$

Daher kann keine Fehlerkomponente von einem Zeitschritt zum nächsten wachsen, solange

$$|e^{a\Delta t}| = \left| \frac{\varepsilon_j^{n+1}}{\varepsilon_j^n} \right| \leq 1$$

erfüllt ist. Der Ausdruck  $\frac{\varepsilon_j^{n+1}}{\varepsilon_j^n}$  heißt *Verstärkungsfaktor* und wird mit  $G$  bezeichnet. Gilt

$$|G| \leq 1,$$

ist das zugrundeliegende Differenzenverfahren stabil.

Da der Einfluß von Randbedingungen nicht berücksichtigt wird, setzt die Stabilitätsanalyse periodische Randbedingungen voraus.

Mit Hilfe der Exponentialdarstellung komplexer Zahlen läßt sich der Verstärkungsfaktor auch in der Form

$$G = |G|e^{i\phi}$$

darstellen. Dabei ist  $\phi$  der Phasenwinkel, der durch

$$\phi = \tan^{-1} \frac{\operatorname{Im}(G)}{\operatorname{Re}(G)}$$

gegeben ist. Damit ergibt die Stabilitätsanalyse nicht nur den Verstärkungsfaktor eines Differenzenverfahrens, sondern ermöglicht auch die Berechnung der zugehörigen Phase. Eine bessere Anschauung als die Phase liefert der *relative Phasenfehler*, der durch den Quotienten  $\phi/\phi_e$  beschrieben wird, wobei  $\phi_e$  den Phasenwinkel der exakten Lösung der zugrundeliegenden Gleichung bezeichnet. Für seine Berechnung ist zunächst der Verstärkungsfaktor der exakten Lösung zu bestimmen.

Betrachten wir die eindimensionale lineare Transportgleichung

$$u_t + cu_x = 0, \quad c > 0, \quad (4.28)$$

die für das Anfangswertproblem

$$u(x, 0) = F(x), \quad -\infty < x < \infty,$$

die exakte Lösung

$$u(x, t) = F(x - ct)$$

besitzt, erhalten wir den exakten Verstärkungsfaktor dieser Gleichung, indem wir den Ansatz

$$u = e^{at} e^{ik_m x}$$

in (4.28) einsetzen. Es ergibt sich  $a = -ik_m c$  und damit

$$u = e^{ik_m(x-ct)},$$

so daß der exakte Verstärkungsfaktor durch

$$G_e = \frac{u(t + \Delta t)}{u(t)} = \frac{e^{ik_m[x-c(t+\Delta t)]}}{e^{ik_m(x-ct)}} = e^{-ik_m c \Delta t}$$

gegeben ist. Da darüber hinaus

$$G_e = e^{i\phi_e}$$

gilt, ergibt sich für den Phasenfehler  $\phi_e$  der exakten Gleichung

$$\phi_e = -k_m c \Delta t = -\beta \nu,$$

wobei  $\nu = \frac{c\Delta t}{\Delta x}$  die CFL-Zahl bezeichnet. Damit ist der Phasenfehler der exakten Gleichung berechnet, und der relative Phasenfehler eines Finite-Differenzen-Verfahrens für die lineare Transportgleichung (4.28) kann angegeben werden.

### Von Neumann Analyse für das aus Filterschritt und Lax-Wendroff-Schema zusammengesetzte Verfahren

Im Folgenden wird die von Neumannsche Stabilitätsanalyse für das aus Lax-Wendroff-Verfahren und Filter (4.25) zusammengesetzte Differenzenschema durchgeführt. Die zugrundeliegende Differentialgleichung sei die lineare Transportgleichung (4.28).

Da wir eine Iteration der Filtergleichung nach jedem Schritt des Lax-Wendroff-Verfahrens

durchführen, setzen wir diese beiden Teilschritte zunächst zu einer Differenzengleichung zusammen. Für die lineare Transportgleichung (4.28) mit konstantem Koeffizienten  $c > 0$  ist das Lax–Wendroff–Verfahren durch

$$\tilde{U}_j^n = U_j^n - \frac{c\Delta t}{2\Delta x}(U_{j+1}^n - U_{j-1}^n) + \frac{c^2\Delta t^2}{2\Delta x^2}(U_{j+1}^n - 2U_j^n + U_{j-1}^n)$$

gegeben. Es liefert ein Zwischenergebnis  $\tilde{U}^n$ , das in die Filtergleichung

$$U_j^{n+1} = \tilde{U}_j^n + \frac{\Delta t}{\Delta x}d_j(\tilde{U}_{j+1}^n - 2\tilde{U}_j^n + \tilde{U}_{j-1}^n)$$

eingesetzt wird. Mit der Bezeichnung  $\nu = \frac{c\Delta t}{\Delta x}$  für die CFL–Zahl ergibt sich insgesamt

$$\begin{aligned} U_j^{n+1} = U_j^n & - \frac{\nu}{2}(U_{j+1}^n - U_{j-1}^n) + \frac{\nu^2}{2}(U_{j+1}^n - 2U_j^n + U_{j-1}^n) \\ & + \frac{\nu}{c}d_j\left[U_{j+1}^n - \frac{\nu}{2}(U_{j+2}^n - U_j^n) + \frac{\nu^2}{2}(U_{j+2}^n - 2U_{j+1}^n + U_j^n) \right. \\ & - 2(U_j^n - \frac{\nu}{2}(U_{j+1}^n - U_{j-1}^n) + \frac{\nu^2}{2}(U_{j+1}^n - 2U_j^n + U_{j-1}^n)) \\ & \left. + U_{j-1}^n - \frac{\nu}{2}(U_j^n - U_{j-2}^n) + \frac{\nu^2}{2}(U_j^n - 2U_{j-1}^n + U_{j-2}^n)\right]. \end{aligned}$$

Setzen wir den Ansatz

$$U_j^n = e^{at}e^{ik_mx}$$

in diese Gleichung ein, wobei wir zur Vereinfachung der Schreibweise den Index  $m$  von  $k_m$  vernachlässigen, erhalten wir

$$\begin{aligned} e^{a\Delta t} &= 1 - \frac{\nu}{2}(e^{ik\Delta x} - e^{-ik\Delta x}) + \frac{\nu^2}{2}(e^{ik\Delta x} - 2 + e^{-ik\Delta x}) \\ &+ \frac{\nu}{c}d_j\left[e^{ik\Delta x} - \frac{\nu}{2}(e^{2ik\Delta x} - 1) + \frac{\nu^2}{2}(e^{2ik\Delta x} - 2e^{ik\Delta x} + 1) \right. \\ &- 2(1 - \frac{\nu}{2}(e^{ik\Delta x} - e^{-ik\Delta x}) + \frac{\nu^2}{2}(e^{ik\Delta x} - 2 + e^{-ik\Delta x})) \\ &\left. + e^{-ik\Delta x} - \frac{\nu}{2}(1 - e^{-2ik\Delta x}) + \frac{\nu^2}{2}(1 - 2e^{-ik\Delta x} + e^{-2ik\Delta x})\right] \end{aligned}$$

und damit

$$\begin{aligned} G &= 1 - i\nu \sin \beta + \frac{\nu^2}{2}(2 \cos \beta - 2) + \frac{\nu}{c}d_j\left[2 \cos \beta - \frac{\nu}{2}(2i \sin(2\beta) - 4i \sin \beta) \right. \\ &\quad \left. + \frac{\nu^2}{2}(6 + 2 \cos(2\beta) - 4 \cos \beta - 4 \cos \beta) - 2\right] \\ &= 1 + \nu^2(\cos \beta - 1) + d_j\frac{\nu}{c}(2 \cos \beta + \frac{\nu^2}{2}(2 \cos(2\beta) - 8 \cos \beta + 6) - 2) \\ &\quad - i\nu \sin \beta(1 + 2d_j\frac{\nu}{c}(\cos \beta - 1)). \end{aligned}$$

Da der Faktor  $d_j$  die Werte Null und eins annehmen kann, sind zwei Fälle zu unterscheiden:

**1. Fall**  $d_j = 0$ : Für diese Situation ergibt sich

$$G = 1 - \nu^2(1 - \cos \beta) - i\nu \sin \beta,$$

so daß der Verstärkungsfaktor  $|G|$  durch

$$|G| = \sqrt{(1 - \nu^2(1 - \cos \beta))^2 + \nu^2 \sin^2 \beta} \quad (4.29)$$

gegeben ist und für den Phasenwinkel

$$\phi = \tan^{-1} \frac{\operatorname{Im}(G)}{\operatorname{Re}(G)} = \tan^{-1} \frac{-\nu \sin \beta}{1 - \nu^2(1 - \cos \beta)}$$

gilt. Da die Filterroutine für  $d_j = 0$  inaktiv ist und das zusammengesetzte Verfahren in diesem Fall lediglich aus dem Lax-Wendroff-Schritt besteht, stimmen Amplitude und Phase mit denen des Lax-Wendroff-Verfahrens überein.

**2. Fall**  $d_j = 1$ : Für diese Situation wird obige Gleichung zu

$$\begin{aligned} G = & 1 - \nu^2(1 - \cos \beta) + \frac{\nu}{c}(2 \cos \beta - 2 + \frac{\nu^2}{2}(\cos(2\beta) - 8 \cos \beta + 6)) \\ & - i\nu \sin \beta(1 + \frac{2\nu}{c}(\cos \beta - 1)), \end{aligned}$$

so daß der Verstärkungsfaktor durch

$$\begin{aligned} |G| = & \left[ (1 - \nu^2(1 - \cos \beta) + \frac{\nu}{c}(2 \cos \beta - 2 + \frac{\nu^2}{2}(2 \cos(2\beta) \right. \\ & \left. - 8 \cos \beta + 6)))^2 + (\nu \sin \beta(1 + \frac{2\nu}{c}(\cos \beta - 1)))^2 \right]^{1/2} \end{aligned} \quad (4.30)$$

und die Phase durch

$$\phi = \tan^{-1} \frac{-\nu \sin \beta(1 + \frac{2\nu}{c}(\cos \beta - 1))}{1 - \nu^2(1 - \cos \beta) + \frac{\nu}{c}(2 \cos \beta - 2 + \frac{\nu^2}{2}(2 \cos(2\beta) - 8 \cos \beta + 6))} \quad (4.31)$$

gegeben ist.

Zur Analyse des Verstärkungsfaktors und der Phase des zusammengesetzten Differenzenverfahrens betrachten wir im Folgenden die polare und kartesische Darstellung von  $|G|$  und  $\phi$  in Abhängigkeit des Parameters  $\beta$ . Abbildung 4.26 zeigt den Verstärkungsfaktor für das reine Lax-Wendroff-Verfahren in Polarkoordinatendarstellung, das heißt, dargestellt ist die Abbildung

$$\vec{r} = \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} |G| \cos \beta \\ |G| \sin \beta \end{pmatrix}$$

mit  $\beta \in [0, \pi]$ .

Entsprechend ist in Abbildung 4.27 der Verstärkungsfaktor für das aus Lax-Wendroff- und Filterschritt zusammengesetzte Verfahren für  $c = 1$  und  $c = 10$  dargestellt. Die Abbildungen 4.28 und 4.29 zeigen die entsprechenden kartesischen Darstellungen, in denen  $|G|$  gegen  $\beta$  aufgetragen ist.

Bei der Darstellung der Phase  $\phi$  muß darauf geachtet werden, daß beim Auflösen der Gleichung

$$\tan \phi = \frac{\operatorname{Im}(G)}{\operatorname{Re}(G)}$$

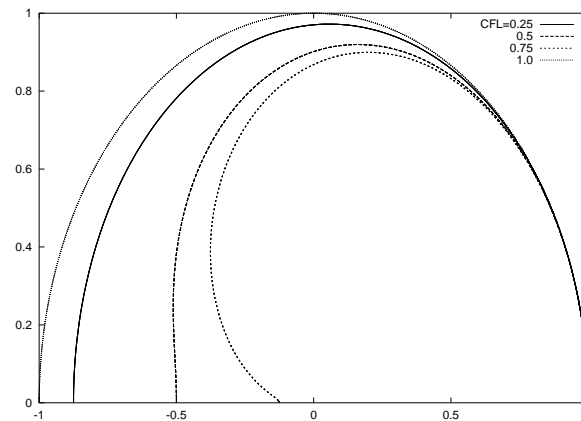


Abbildung 4.26: Verstärkungsfaktor des Lax-Wendroff-Verfahrens für verschiedene CFL-Zahlen in Polarkoordinatendarstellung

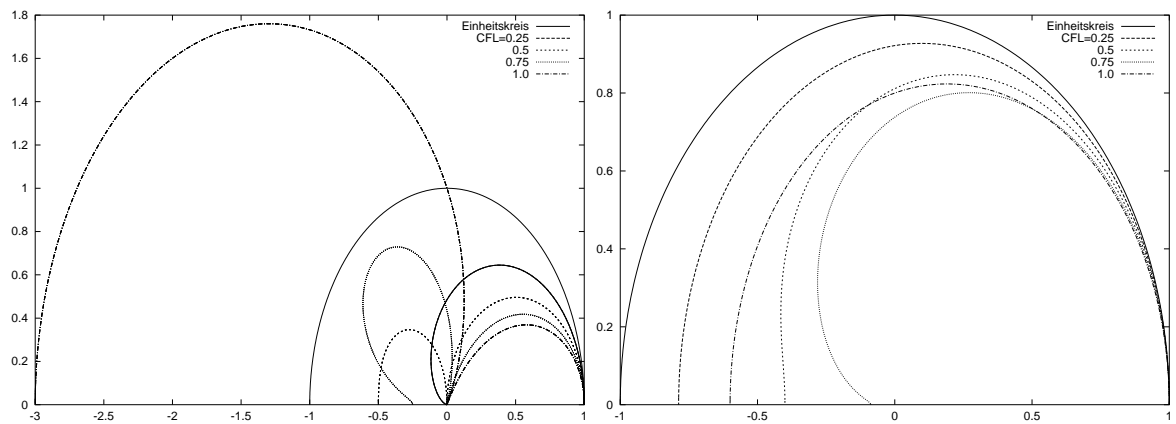


Abbildung 4.27: Verstärkungsfaktor des zusammengesetzten Verfahrens für  $c = 1$  (links) und  $c = 10$  (rechts) und verschiedene CFL-Zahlen in Polarkoordinatendarstellung

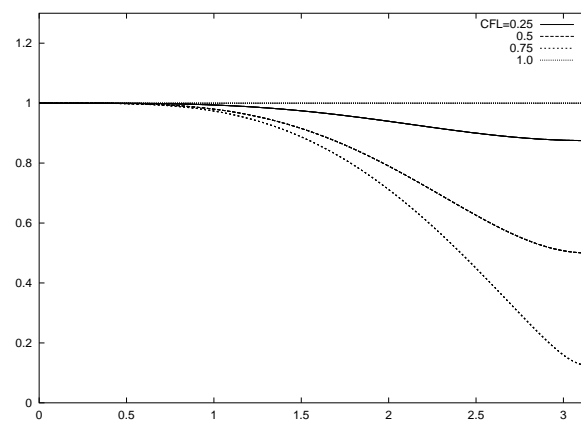


Abbildung 4.28: Verstärkungsfaktor des Lax-Wendroff-Verfahrens für verschiedene CFL-Zahlen in kartesischer Darstellung

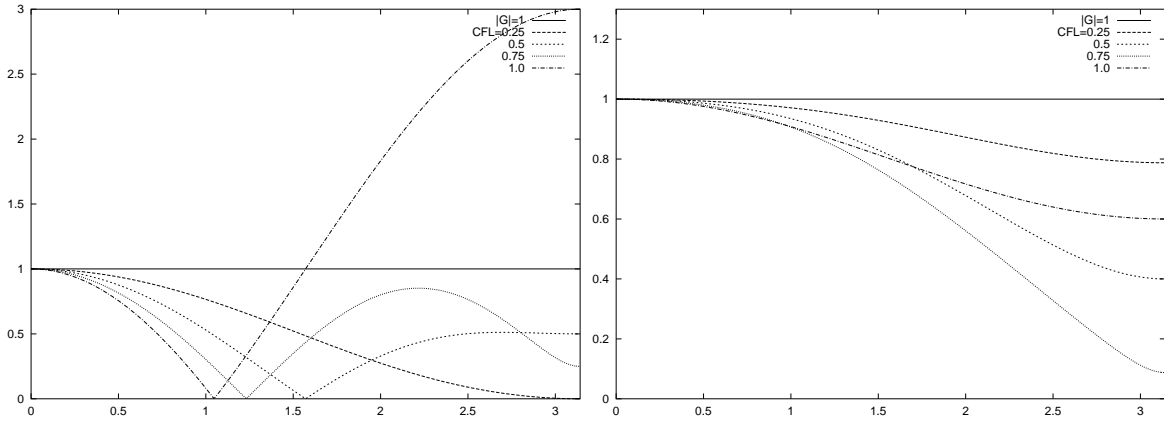


Abbildung 4.29: Verstärkungsfaktor des zusammengesetzten Verfahrens für  $c = 1$  (links) und  $c = 10$  (rechts) und verschiedene CFL-Zahlen in kartesischer Darstellung

der richtige Bereich für  $\phi$  verwendet wird. Für  $\varphi \in (-\pi, \pi)$  ergibt sich die allgemeine Lösung der Gleichung

$$\tan \varphi = \frac{y}{x}$$

aus

$$\varphi = \begin{cases} \arctan \frac{y}{x} & : x > 0, y \geq 0 \\ \frac{\pi}{2} & : x = 0, y > 0 \\ \arctan \frac{y}{x} + \pi & : x < 0, y \geq 0 \\ \arctan \frac{y}{x} - \pi & : x < 0, y < 0 \\ \arctan \frac{y}{x} & : x > 0, y < 0 \end{cases}, \quad (4.32)$$

so daß diese Rechenvorschrift mit  $x = \operatorname{Re}(G)$  und  $y = \operatorname{Im}(G)$  die richtige Auflösung von  $\tan \phi = \operatorname{Im}(G)/\operatorname{Re}(G)$  nach  $\phi$  liefert.

Für eine bessere Anschauung ist es vorteilhaft, nicht die Phase  $\phi$  sondern die relative Phase  $\phi/\phi_e$  des Differenzenverfahrens darzustellen, wobei  $\phi_e$  die Phase der exakten Lösung bezeichnet, die, wie oben hergeleitet, durch

$$\phi_e = -\beta\nu$$

gegeben ist.

Die Polarkoordinatendarstellung der relativen Phase ergibt sich aus der Abbildung

$$\vec{r} = \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \frac{\phi}{\phi_e} \cos \beta \\ \frac{\phi}{\phi_e} \sin \beta \end{pmatrix}$$

und ist in Bild 4.30 für das Lax-Wendroff-Verfahren und in Bild 4.31 für das zusammengesetzte Verfahren mit integriertem Filter für  $c = 1$  und  $c = 10$  dargestellt. Die Abbildungen 4.32 und 4.33 zeigen die entsprechenden kartesischen Darstellungen der relativen Phase des Lax-Wendroff-Schemas und des zusammengesetzten Verfahrens, wobei jeweils  $\phi/\phi_e$  gegen  $\beta$  aufgetragen ist.

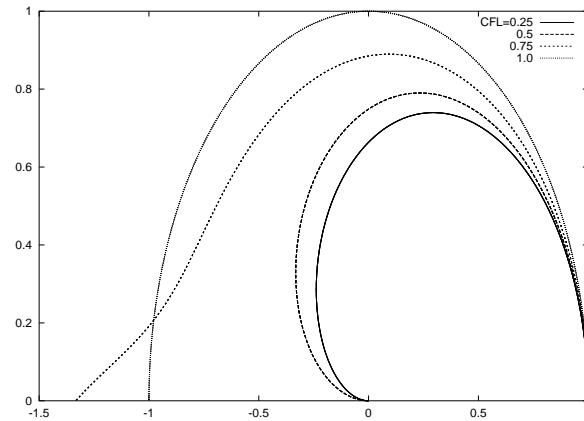


Abbildung 4.30: Relative Phase des Lax-Wendroff-Verfahrens für verschiedene CFL-Zahlen in Polarkoordinaten

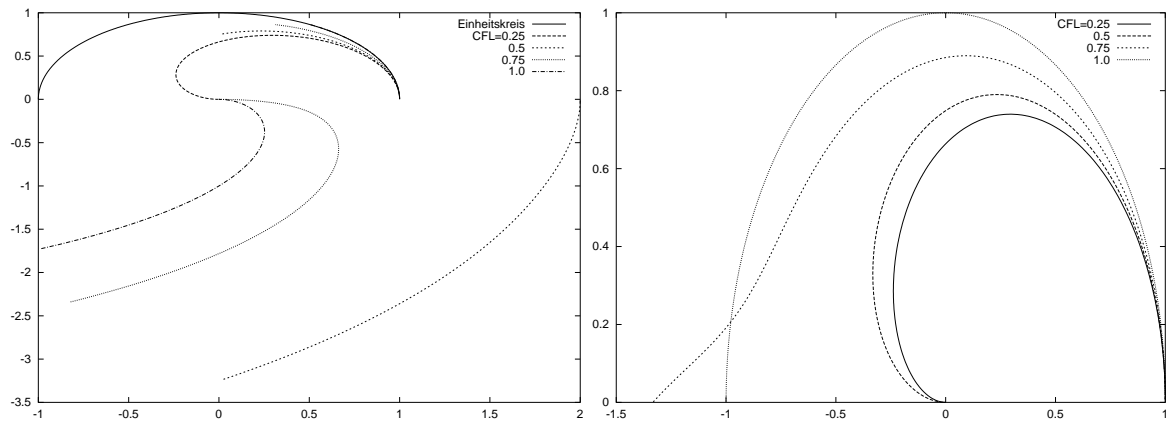


Abbildung 4.31: Relative Phase des zusammengesetzten Verfahrens für  $c = 1$  (links) und  $c = 10$  (rechts) und verschiedene CFL-Zahlen in Polarkoordinaten

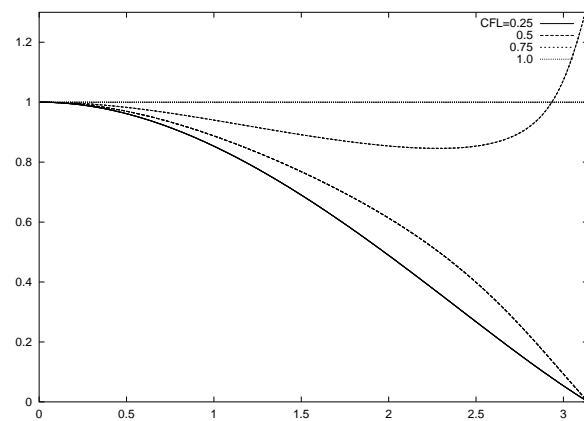


Abbildung 4.32: Kartesische Darstellung der relativen Phase des Lax-Wendroff-Verfahrens für verschiedene CFL-Zahlen



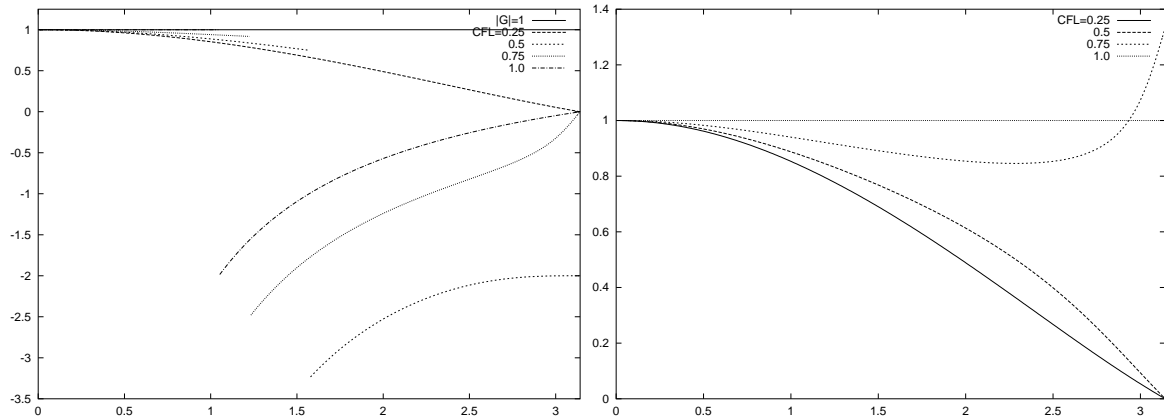


Abbildung 4.33: Kartesische Darstellung der relativen Phase des zusammengesetzten Verfahrens für  $c = 1$  (links) und  $c = 10$  (rechts) und verschiedene CFL-Zahlen

## Interpretation der Ergebnisse

### Bedeutung der Darstellungen für den Verstärkungsfaktor

Die Schaubilder 4.26 und 4.28 des Verstärkungsfaktors  $|G|$  für das reine Lax–Wendroff–Verfahren zeigen, daß für jede dargestellte CFL–Zahl  $|G| \leq 1$  gilt und das Verfahren damit stets stabil ist. In Abbildung 4.26 erkennt man dies daran, daß keine Kurve außerhalb des Einheitshalbkreises liegt, und Abbildung 4.28 verdeutlicht das Ergebnis, da keine Kurve oberhalb der Geraden  $|G| = 1$  verläuft. Eine Abschätzung der Gleichung (4.29),

$$|G| = \sqrt{(1 - \nu^2(1 - \cos \beta))^2 + \nu^2 \sin^2 \beta},$$

zeigt, daß das Lax–Wendroff–Verfahren für alle CFL–Zahlen  $\nu \leq 1$  stabil ist.

Anders sieht es für das aus Lax–Wendroff– und Filterschritt zusammengesetzte Verfahren aus. Die Abbildungen 4.27 und 4.29 zeigen zunächst eine starke Abhängigkeit des Verstärkungsfaktors von dem Koeffizienten  $c$  der Transportgleichung. Für  $c = 1$  liegen die dargestellten Kurven bis einschließlich für die CLF–Zahl  $\nu = 0.75$  innerhalb des Einheitshalbkreises (Abbildung 4.27) beziehungsweise unterhalb der Geraden  $|G| = 1$  (Abbildung 4.29). Für die CFL–Zahl  $\nu = 1$  ist der Verstärkungsfaktor ab einer gewissen Frequenz größer als eins, so daß das resultierende Verfahren instabil wird. Dies gilt nicht nur für  $\nu = 1$ , schon ab CFL–Zahlen von 0.8 ist der Verstärkungsfaktor teilweise größer als eins. Im Gegensatz zum Lax–Wendroff–Verfahren, für das die Verstärkungsfaktoren jeweils monoton fallende Funktionen von  $\beta$  sind, nehmen die Verstärkungsfaktoren der Filtermethode für  $\nu > 0.25$  mit wachsendem  $\beta$  zunächst ab, erreichen das Minimum  $|G| = 0$  und steigen dann wieder an. Für  $c = 10$  liegen alle dargestellten Verstärkungsfaktoren innerhalb des Einheitshalbkreises beziehungsweise unterhalb der Geraden  $|G| = 1$ , so daß das Filterverfahren stets stabil ist. Darüber hinaus ähneln die Kurvenverläufe der dargestellten Verstärkungsfaktoren für  $c = 10$  denen des Lax–Wendroff–Verfahrens erheblich, wobei die Filtermethode allerdings für jede CFL–Zahl stärker gedämpfte Amplituden aufweist. Ebenso wie beim Lax–Wendroff–Verfahren nimmt die Dämpfung mit steigender Frequenz zu. Die Ähnlichkeiten zwischen den Verstärkungsfaktoren des zusammengesetzten Verfahrens und dem Lax–Wendroff–Verfahren

sind für größer werdendes  $c$  bereits an Gleichung (4.30) ersichtlich. Die zusätzlichen Terme in der Gleichung für den Verstärkungsfaktor des gefilterten Verfahrens sind alle mit dem Faktor  $\frac{\nu}{c}$  multipliziert, so daß sie mit wachsendem  $c$  immer kleiner werden und die Routine in das reine Lax–Wendroff–Verfahren übergeht.

Die theoretisch hergeleiteten Stabilitätseigenschaften decken sich exakt mit den Beobachtungen bei der numerischen Simulation, wobei wir den Schalter  $d_j$  der Filtergleichung (4.25) konstant auf eins setzen, damit wir die in den Abbildungen 4.27 und 4.29 dargestellte Situation exakt simulieren. Das numerische Lösen der linearen Transportgleichung mit einer Unstetigkeit als Anfangsbedingung ergibt für  $c = 1$  und CFL–Zahlen  $\nu < 0.8$  ein stabiles Verfahren, für größere CFL–Zahlen ist es instabil. Dagegen bleibt die Stabilität des Verfahrens für  $c = 10$  selbst bei  $\nu = 1$  erhalten.

### Bedeutung der Darstellungen für die relative Phase

Die Darstellungen der relativen Phase sind wie folgt zu interpretieren. Ist der relative Phasenfehler für ein gegebenes  $\beta$  größer als eins, so besitzt die zugehörige Fourierkomponente der numerischen Lösung eine größere Wellengeschwindigkeit als die der exakten Lösung, und es handelt sich um einen vorausseilenden Phasenfehler. Ist der relative Phasenfehler kleiner als eins, so ist die Wellengeschwindigkeit der numerischen Lösung kleiner als die exakte Wellengeschwindigkeit, und es liegt ein hinterhereilender Phasenfehler vor.

Abbildungen 4.30 und 4.32 zeigen, daß der relative Phasenfehler des Lax–Wendroff–Verfahrens überwiegend kleiner eins ist, was auf einen hinterhereilenden Phasenfehler hinweist. Nur für große Wellenzahlen, also hohe Frequenzen, und  $\sqrt{0.5} < \nu < 1$  tritt ein vorausseilender Phasenfehler auf.

Für das aus Lax–Wendroff–Schritt und Filter zusammengesetzte Verfahren zeigen die Abbildung 4.31 und 4.33, daß der relative Phasenfehler für  $c = 1$  stets kleiner oder gleich eins ist, so daß das Verfahren in diesem Fall für alle Frequenzen einen hinterhereilenden Phasenfehler aufweist. Auffällig ist, daß die Kurven der relativen Phasen für alle CFL–Zahlen  $\nu > 0.25$  einen deutlichen Sprung im Bereich  $1 < \beta < 1.6$  aufzeigen. Für  $c = 10$  existiert ein solcher Sprung nicht und die dargestellten relativen Phasen sind, ebenso wie die Amplituden, denen des reinen Lax–Wendroff–Verfahrens sehr ähnlich.

### Unterschiede der Verfahren für CFL= 0.1

Um die Unterschiede zwischen dem neu entwickelten Verfahren und der herkömmlichen Lax–Wendroff–Methode genauer zu untersuchen, vergleichen wir im Folgenden Amplitude und Phase beider Verfahren für die CFL–Zahl 0.1.

Die Abbildungen 4.34 und 4.35 zeigen die jeweiligen Verstärkungsfaktoren der Verfahren für  $c = 1$  und  $c = 10$  in polarer und kartesischer Darstellung, in den Abbildungen 4.36 und 4.37 sind die relativen Phasenfehler in entsprechender Weise dargestellt.

Für  $c = 1$  ist der Verstärkungsfaktor des zusammengesetzten Verfahrens besonders für hohe Frequenzen deutlich kleiner als der des reinen Lax–Wendroff–Verfahrens, was auf eine deutlich stärkere Dämpfung der Amplituden bei der Filterimplementierung hinweist. Lediglich für sehr niedrige Frequenzen stimmen die Amplituden beider Verfahren überein. Auch für  $c = 10$  nimmt der Unterschied zwischen den Amplituden für wachsende Frequenz zu und das

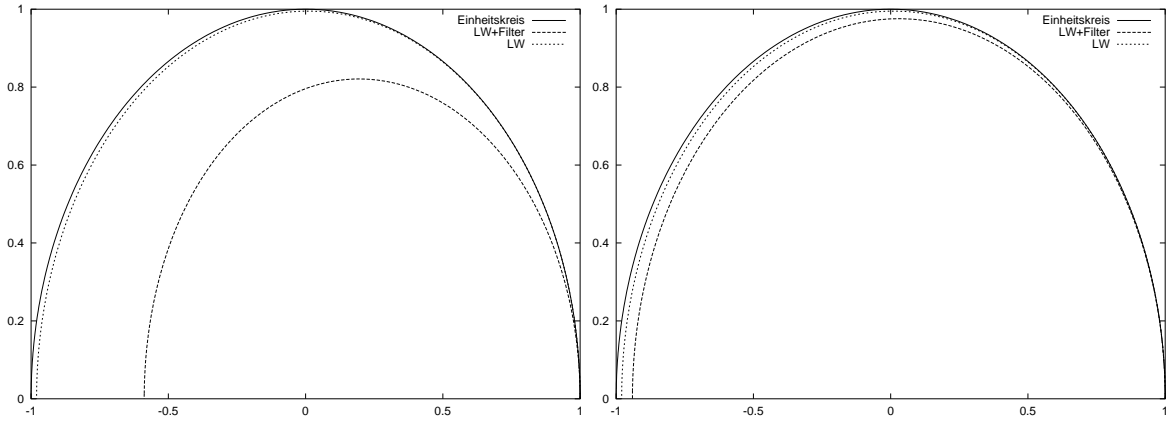


Abbildung 4.34: Verstärkungsfaktor für Lax–Wendroff– und zusammengesetztes Verfahren für CFL= 0.1 und  $c = 1$  (links) beziehungsweise  $c = 10$  (rechts) in Polarkoordinaten

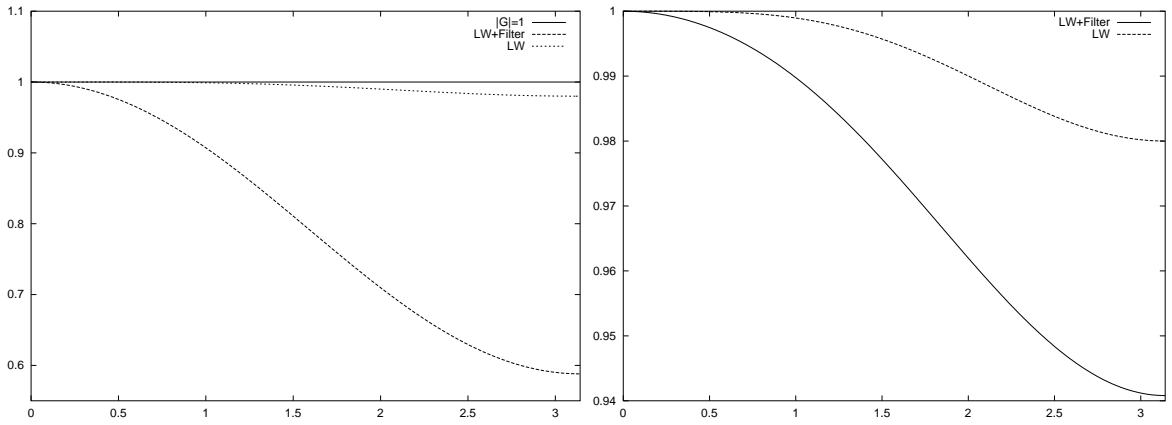


Abbildung 4.35: Verstärkungsfaktor für Lax–Wendroff– und zusammengesetztes Verfahren für CFL= 0.1 und  $c = 1$  (links) beziehungsweise  $c = 10$  (rechts) in kartesischen Koordinaten

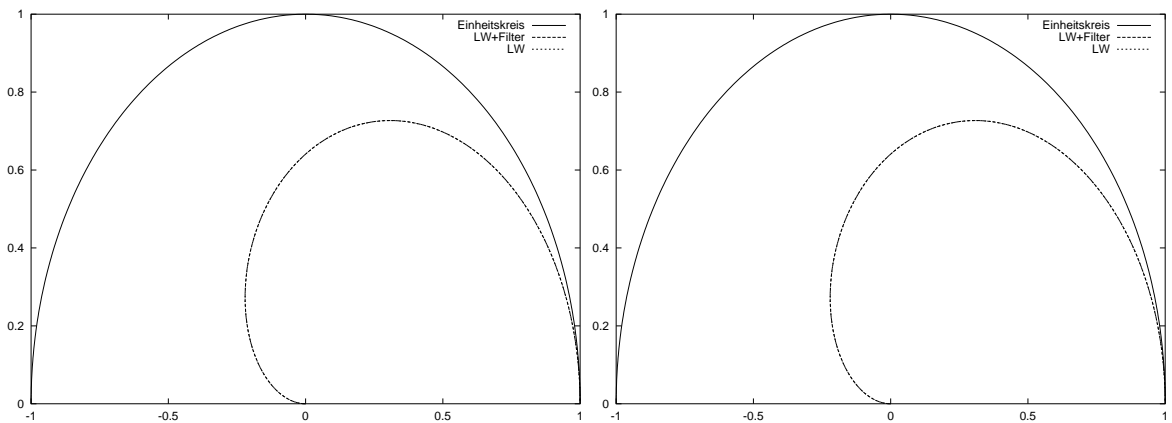


Abbildung 4.36: Relative Phase des Lax–Wendroff– und zusammengesetzten Verfahrens für CFL= 0.1 und  $c = 1$  (links) beziehungsweise  $c = 10$  (rechts) in Polarkoordinaten

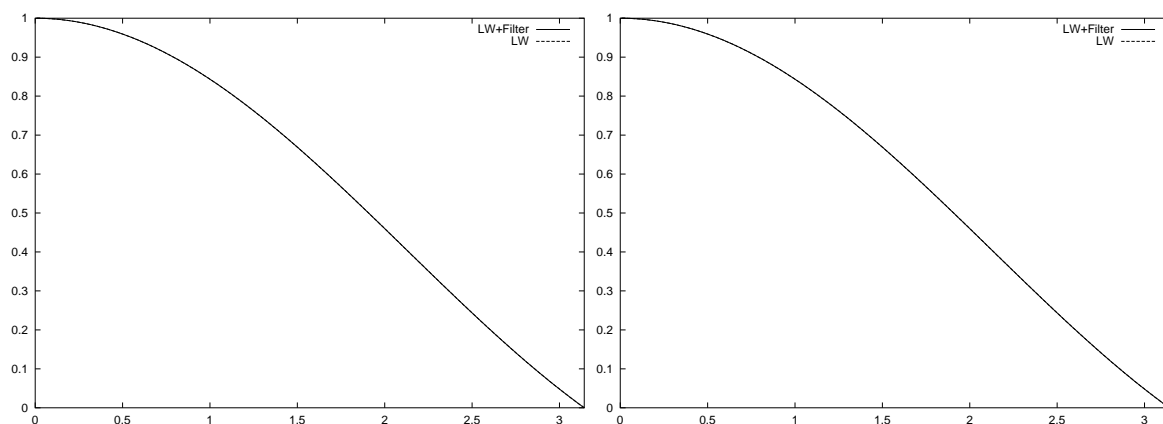


Abbildung 4.37: Relative Phase des Lax–Wendroff– und zusammengesetzten Verfahrens für  $CFL= 0.1$  und  $c = 1$  (links) beziehungsweise  $c = 10$  (rechts) in kartesischen Koordinaten

gefilterte Verfahren besitzt einen kleineren Verstärkungsfaktor als die Lax–Wendroff–Methode, allerdings sind die Abweichungen beider Verfahren in diesem Fall weitaus kleiner. Die relative Phase des gefilterten und ungefilterten Lax–Wendroff–Verfahrens stimmen sowohl für  $c = 1$  als auch für  $c = 10$  exakt überein.

### Zusammenhang zu den numerischen Simulationen

Um die theoretisch gewonnenen Ergebnisse mit Hilfe numerischer Untersuchungen zu überprüfen, betrachten wir im Folgenden den linearen Transport einer Unstetigkeit und eines Wellenpakets. Da die Fouriertransformierte einer Unstetigkeit hauptsächlich aus hochfrequenten Anteilen besteht, spiegelt die numerische Simulation ihres linearen Transportes die Situation hoher Frequenzen wieder. Der Transport eines Wellenpaketes ermöglicht dagegen, feste Wellenzahlen und damit bestimmte Werte für  $\beta$  vorzugeben, für die die theoretischen Aussagen mit der Numerik verglichen werden können.

#### Linearer Transport einer Unstetigkeit

Zunächst betrachten wir den linearen Transport einer Unstetigkeit. Abbildung 4.38 vergleicht die numerischen Ergebnisse, die wir mit Hilfe des zusammengesetzten Verfahrens und der reinen Lax–Wendroff–Methode erhalten. Um der theoretisch untersuchten Situation genau zu entsprechen, ist der Koeffizient  $d_j$  der Filterroutine dabei stets eins, so daß der Filter permanent aktiviert ist. Im linken Teil der Abbildung sind die Resultate für  $c = 1$ , im rechten für  $c = 10$  dargestellt. Zusätzlich ist jeweils die exakte Lösung des linearen Transportproblems angegeben. In Übereinstimmung mit den theoretischen Ergebnissen für hohe Frequenzen ist zu erkennen, daß das gefilterte Verfahren eine stärkere Amplitudendämpfung als das Lax–Wendroff–Verfahren aufweist, wobei die Dämpfung für  $c = 1$  stärker ist als für  $c = 10$ .

Der relative Phasenfehler des Lax–Wendroff–Verfahrens schlägt sich in der numerischen Simulation dadurch nieder, daß sich die für dieses Verfahren typischen Oszillationen ausbilden. Da der Phasenfehler, wie die Theorie zeigt, hinterhereilenden Charakter hat, entstehen diese Oszillationen bei positiver Ausbreitungsgeschwindigkeit links von der Unstetigkeit und eilen

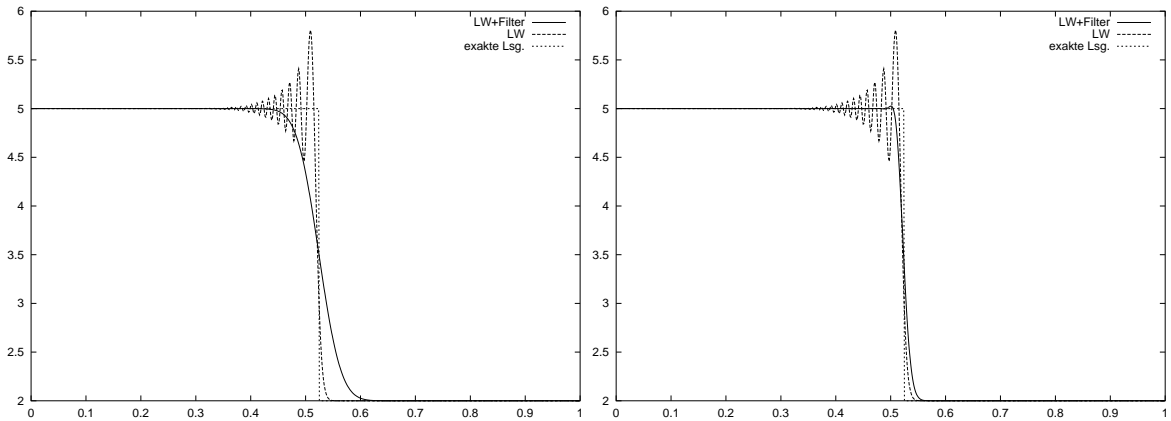


Abbildung 4.38: Vergleich zwischen exakter, gefilterter und Lax-Wendroff-Lösung der linearen Transportgleichung für  $c = 1$  (links) und  $c = 10$  (rechts) bei permanenter Filteraktivierung

dem Stoß damit hinterher. Der relative Phasenfehler des zusammengesetzten Filterverfahrens stimmt exakt mit dem der Lax-Wendroff-Lösung überein, so daß auch die gefilterte Routine Oszillationen links vom Stoß erzeugt. Da die Amplituden bei der Filteraktivierung allerdings stark gedämpft werden, sind diese Oszillationen für  $c = 10$  zum größten Teil und für  $c = 1$  vollständig geglättet und daher in Abbildung 4.38 kaum oder gar nicht zu erkennen. Dennoch ist für den linearen Transport einer Unstetigkeit, dessen Fouriertransformierte hauptsächlich aus hochfrequenten Anteilen besteht, die Analogie zwischen theoretischen und numerischen Ergebnissen gezeigt.

Die zu starke Glättungseigenschaft der zusammengesetzten Verfahrens, die in Abbildung 4.38 besonders für  $c = 1$  auffällt, ist darauf zurückzuführen, daß die Filterroutine für die vorangegangenen Experimente für jeden Gitterpunkt ausgeführt wurde. Aktivieren wir den Filterschritt dagegen, wie ursprünglich konstruiert, nur für Extremwerte, erhalten wir, statt den Lösungen in Abbildung 4.38, die in Abbildung 4.39 dargestellten Ergebnisse. Für den

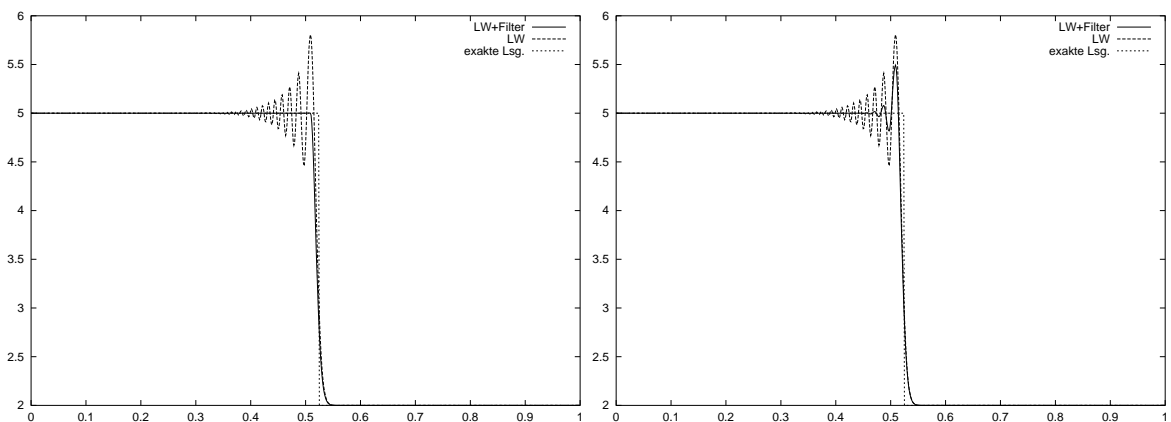


Abbildung 4.39: Vergleich zwischen exakter, gefilterter und Lax-Wendroff-Lösung der linearen Transportgleichung für  $c = 1$  (links) und  $c = 10$  (rechts), wobei der Filter nur an Extremwerten aktiviert ist

Fall  $c = 1$  ist die Steilheit der Unstetigkeit fast vollständig erhalten und die Oszillationen des Lax–Wendroff–Verfahrens sind geglättet, so daß das Filterverfahren sehr gute Resultate liefert. Auch für  $c = 10$  ist die Stoßhöhe erhalten, allerdings treten noch Restoszillationen vor der Unstetigkeit auf.

### Linearer Transport eines Wellenpaketes

Um die theoretisch gewonnenen Aussagen über Amplitude und Phase auch für niederfrequente Signale mit Hilfe numerischer Simulationen zu überprüfen, betrachten wir im Folgenden ein sinusförmiges, niederfrequentes Wellenpaket, das es uns ermöglicht, die numerischen Ergebnisse der linearen Transportgleichung für bestimmte Frequenzen beziehungsweise Winkel  $\beta$  auszuwerten.

Wir betrachten zunächst die numerische Lösung der linearen Transportgleichung mit Hilfe des Lax–Wendroff–Verfahrens für das niederfrequente Signal

$$y(x, 0) = \sin\left(\frac{\beta}{\Delta x} \cdot x\right) = \sin\left(\frac{1}{10}\pi \cdot \frac{x}{\Delta x}\right).$$

Bei einer Ausbreitungsgeschwindigkeit von  $c = 1$  erhalten wir für die CFL–Zahl  $\nu = 0.1$  nach der Zeit  $t = 2.5$  die im linken Teil von Abbildung 4.40 dargestellte Lösung. Da das

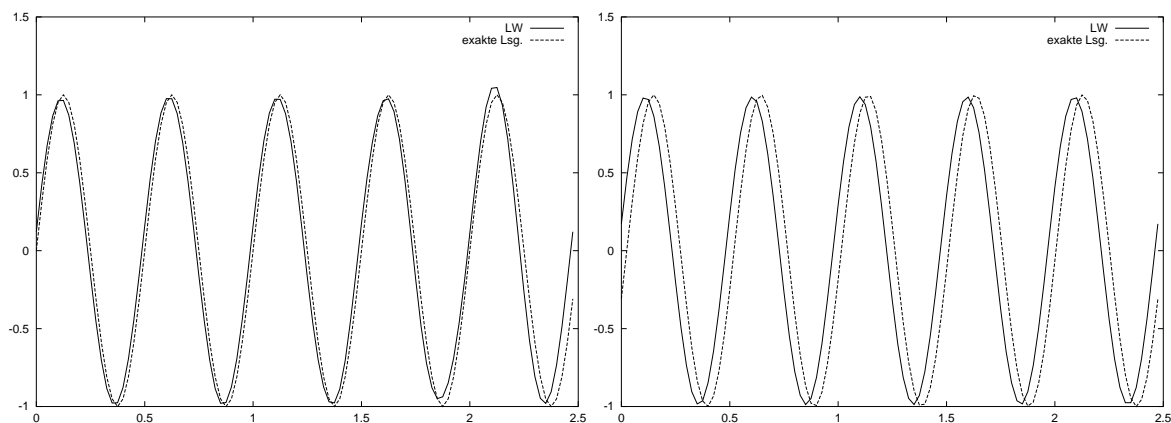


Abbildung 4.40: Exakte Lösung und Lax–Wendroff–Lösung der linearen Transportgleichung für ein niederfrequentes Signal mit  $\beta = 0.1\pi$  (links) und  $\beta = 0.101\pi$  (rechts)

Lax–Wendroff–Verfahren einen hinterhereilenden Phasenfehler besitzt, ist eine leichte Verschiebung der zugehörigen Lösung nach links gegenüber der exakten Lösung zu erkennen. Für kleine Frequenzen und damit kleine Winkel  $\beta$  ist der Phasenfehler des Verfahrens sehr klein, so daß die Abweichungen nur minimal sind. Wie Abbildung 4.37 links zeigt, nimmt der Phasenfehler für wachsendes  $\beta$  stark zu, was sich in den numerischen Experimenten schon für einen minimal größeren Winkel  $\beta = 0.101\pi$  bestätigen läßt, wie der rechte Teil der Abbildung 4.40 zeigt. Die Lax–Wendroff–Lösung ist gegenüber der exakten Lösung weitaus deutlicher nach links verschoben als noch für  $\beta = 0.1\pi$ .

Abbildung 4.41 vergleicht die gefilterte und ungefilterte Version der Lax–Wendroff–Lösung nach der Zeit  $t = 0.2$ , wobei  $\beta = 0.1\pi$  und  $\nu = 0.1$  gilt. Wie die theoretischen Untersuchungen in den Abbildungen 4.36 und 4.37 zeigen, stimmen beide Resultate in ihrer Phase überein.

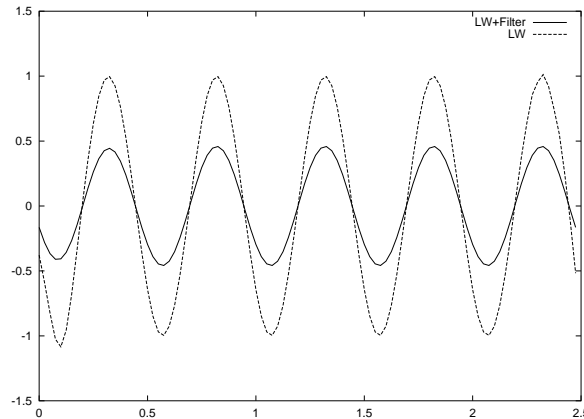


Abbildung 4.41: Vergleich der reinen und gefilterten Lax-Wendroff-Lösung für  $\beta = 0.1\pi$

Da der Verstärkungsfaktor der gefilterten Methode kleiner ist als der des Lax-Wendroff-Verfahrens, sind die Amplituden des gefilterten Signals gedämpft. Diese Dämpfung bewirkt, daß das Signal schon nach kurzer Zeit völlig zerstört wird, und macht numerische Untersuchungen für größere Winkel  $\beta$  oder größere CFL-Zahlen  $\nu$  unmöglich. Dennoch liefern die Fälle, für die eine numerische Analyse möglich ist, Übereinstimmung mit den theoretisch hergeleiteten Eigenschaften der Verfahren.

## Schlußfolgerungen

Die von Neumannsche Stabilitätsanalyse ermöglicht es also, Aussagen über die Amplitude und Phase von Differenzenverfahren für lineare Problemstellungen zu machen. Da für unsere Anwendungen allerdings nichtlineare Modellgleichungen im Vordergrund stehen, ist die Aussagekraft dieser Analysemethode eher gering. Sie liefert zwar die notwendige Bedingung, daß die Stabilität der linearisierten Modellgleichung mit festgehaltenen Koeffizienten eine notwendige Voraussetzung für die Stabilität der zugehörigen nichtlinearen Gleichung ist, aber eine hinreichende Bedingung liefert diese Methodik nicht.

Die anfängliche Vermutung, mit Hilfe der von Neumannschen Stabilitätsanalyse, abgesehen von der Nichtkonservativität, weitere Ursachen für die Verschiebungsfehler, die bei der Filterimplementierung entstehen, aufzeigen zu können, hat sich nicht bestätigt. Die relativen Phasenfehler, die sich mit dieser Methode berechnen lassen, sind als Dispersionsfehler zu interpretieren, die sich bei der numerischen Simulation für hochfrequente Bereiche, also Unstetigkeiten, als Oszillationen vor oder hinter dem Stoß bemerkbar machen, nicht als Verschiebung der Lösung.

Darüber hinaus wurde bei den Untersuchungen ersichtlich, daß Verschiebungsfehler durch die Implementierung von (4.25) nur auftreten, wenn das zugrundegelegte Modell nichtlinear ist. Ein Filter, der beispielsweise für das einfache Cauchy-Problem für Burgers' Gleichung einen Verschiebungsfehler in der Lösung erzeugt, weist bei seiner Anwendung auf die lineare Transportgleichung keinen solchen Fehler auf.

Weitere numerische Experimente zeigen, daß bei der Implementierung von Filtergleichung (4.25) und Anwendung auf das nichtlineare einfache Cauchy-Problem für Burgers' Gleichung

nur dann ein Verschiebungsfehler in der numerischen Lösung entsteht, wenn der diskrete Schalter  $d_j$  aktiviert ist und für Extremwerte den Wert eins annimmt und sonst  $d_j = 0$  gilt. Wird  $d_j$  dagegen konstant auf eins gesetzt, entsteht der Verschiebungsfehler nicht. Dies zeigt, daß Verschiebungsfehler nicht allein auf die Nichtkonservativität der Filterroutinen zurückgeführt werden können. Auch die Nichtlinearität der zugrundeliegenden Problemstellung und die Verwendung vollständig diskreter Schalter bei der Filterimplementierung können Ursache einer Stoßverschiebung sein.

## 4.2 Prinzip der Signalzerlegung

Bisher wurden neue Filterroutinen für das numerische Lösen hyperbolischer Erhaltungsgleichungen direkt auf den zu filternden Datensatz angewendet. Das folgende Unterkapitel stellt dagegen eine neue Art der Implementierung bekannter Filter dar, die darauf beruht, daß das zu filternde Signal vor der Filteranwendung in zwei Teilsignale zerlegt wird. Die Diffusionsterme, die bei den einzelnen Arbeitsschritten des neuen Verfahrens entstehen, können zu einer Gesamtdiffusion zusammengefaßt werden, die dann wiederum als diskrete Filterroutine interpretiert werden kann.

Zunächst betrachten wir die grundlegende Methode dieser Daten- oder Signalzerlegung und gehen auf ihre Anwendung im Kontext nichtlinearer Filter für Differenzenverfahren höherer Ordnung ein. Es folgt die Analyse der verschiedenen Diffusionsterme und ihre Zusammensetzung sowie ein Vergleich dieser neuen Methodik mit dem aus der Bildverarbeitung bekannten Soft Wavelet Shrinkage. Zum Abschluß präsentieren wir numerische Resultate, die die neuen Verfahren für verschiedene Anwendungsbeispiele liefern.

### 4.2.1 Grundlegende Methode

Die grundlegende Idee bei der Daten- oder Signalzerlegung ist es, ein verrauschtes Eingangssignal, das ohne Beeinträchtigung der physikalischen Strukturen geglättet werden soll, zunächst in zwei Signale zu zerlegen und den nichtlinearen Filter erst im Anschluß daran auf eines dieser Teilsignale anzuwenden, anstatt das Eingangssignal direkt zu filtern.

In der Signalverarbeitung wird zu diesem Zweck wie folgt vorgegangen. Das Eingangssignal, das wir mit  $A$  bezeichnen, wird mit einem einfachen Tiefpaßfilter zerlegt. Es ergibt sich ein vom Rauschen befreites aber auch überglättetes Signal  $B$  sowie das Differenzsignal  $A - B$ . Auf dieses Differenzsignal wird der nichtlineare Filter angewendet, der das Signal  $C$  zurückgibt. Anschließend werden die Signale  $B$  und  $C$  addiert, und ihre Summe liefert das gefilterte Eingangssignal, befreit von Oszillationen ohne Verfälschung der tatsächlichen Signalstrukturen.

Bei der Übertragung dieses Vorgehens auf numerische Lösungsverfahren für hyperbolische Erhaltungsgleichungen kann analog vorgegangen werden. Eine oszillierende Lösung eines Differenzenverfahrens höherer Ordnung stellt das Eingangssignal  $A$  dar, das durch ein Verfahren erster Ordnung, beispielsweise dem Lax–Friedrichs–Verfahren, tiefpaßgefiltert wird. Auf das Differenzsignal wird ein nichtlinearer Filter, hier exemplarisch die diskrete Version der neu



entwickelten Filtergleichung

$$u_t = ((1 - \operatorname{sgn} |u_x|)u_x)_x, \quad (4.33)$$

angewendet, und das Ergebnis wird zu der Lösung des Differenzenverfahrens erster Ordnung addiert.

Prinzipiell gibt es auch hier zwei unterschiedliche Möglichkeiten, das Filterverfahren mit Hilfe der Datenzerlegung zu implementieren. Einerseits läßt sich die Methode so einbauen, daß sie nach vollständiger Beendigung des Differenzenverfahrens in Kraft tritt, andererseits läßt sie sich nach jedem Zeitschritt des Differenzenverfahrens anwenden. Vorteil letzterer Möglichkeit ist dabei wieder, daß das zugrundeliegende Differenzenverfahren keine numerisch stabile Lösung hervorbringen muß und damit auf sehr einfache und rechentechnisch günstige Verfahren, wie dem Zentrale-Differenzen-Verfahren zweiter Ordnung, zurückgegriffen werden kann. Für beide Implementierungsmöglichkeiten des Filterns mit vorheriger Datenzerlegung sind in Abschnitt 4.2.4 Anwendungsbeispiele und Ergebnisse gegeben.

Zunächst gehen wir auf die bei der Datenzerlegungsmethode entstehenden Diffusionsterme ein.

## 4.2.2 Diffusionsanalyse

Basierend auf dem Zentrale-Differenzen-Verfahren als zugrundeliegendes Differenzenschema und dem Lax-Friedrichs-Verfahren als Tiefpaßfilter wird die Methode der Datenzerlegung im Folgenden bezüglich der auftretenden Diffusionsterme untersucht. Dabei berechnen wir die Diffusionen der einzeln angewendeten Differenzenverfahren und setzen diese anschließend zu einer gesamten Diffusion zusammen. Mit Hilfe des resultierenden Diffusionsterms läßt sich eine weitere Implementierungsvariante der Datenzerlegungsmethode konstruieren, bei der die Gesamtdiffusion direkt in das zugrundegelegte Differenzenverfahren implementiert wird.

Für die Diffusionsanalyse betrachten wir das eindimensionale Cauchy-Problem (1.49) für Burgers' Gleichung, das zum Zeitpunkt  $t = 1$  einen Stoß ausbildet. Als zugrundeliegendes Differenzenschema verwenden wir das Zentrale-Differenzen-Verfahren zweiter Ordnung, dessen kontinuierliche Gleichung wir mit Hilfe der Taylorentwicklung und den Bezeichnungen  $h \equiv \Delta x$  und  $k \equiv \Delta t$  berechnen:

$$\begin{aligned} U_i^{n+1} &= U_i^n - \frac{k}{h} \left( \frac{1}{2} (f(U_{i+1}^n) + f(U_i^n)) - \frac{1}{2} (f(U_i^n) + f(U_{i-1}^n)) \right) \\ &= U_i^n - \frac{k}{h} \left( \frac{1}{2} f(U_{i+1}^n) - \frac{1}{2} f(U_{i-1}^n) \right) \\ \Leftrightarrow \frac{U_i^{n+1} - U_i^n}{k} &= - \frac{f(U_{i+1}^n) - f(U_{i-1}^n)}{2h} \\ \Rightarrow u_t + \frac{1}{2} k u_{tt} + O(k^2) &= -f(u)_x + O(h^2). \end{aligned}$$

Gleiches Vorgehen für das Lax–Friedrichs–Verfahren, das als Tiefpaßfilter eingesetzt wird, liefert

$$\begin{aligned}
U_i^{n+1} &= U_i^n - \frac{k}{h} \left( \frac{1}{2} (f(U_{i+1}^n) - f(U_{i-1}^n)) - \frac{h}{2k} (U_{i+1}^n - 2U_i^n + U_{i-1}^n) \right) \\
&= U_i^n - \frac{k}{2h} (f(U_{i+1}^n) - f(U_{i-1}^n)) + \frac{1}{2} (U_{i+1}^n - 2U_i^n + U_{i-1}^n) \\
\Leftrightarrow \frac{U_i^{n+1} - U_i^n}{k} &= -\frac{f(U_{i+1}^n) - f(U_{i-1}^n)}{2h} + \frac{U_{i+1}^n - 2U_i^n + U_{i-1}^n}{2k} \\
\Rightarrow u_t + \frac{1}{2}ku_{tt} + O(k^2) &= -f(u)_x + \frac{h^2}{2k}u_{xx} + O(h^2).
\end{aligned}$$

Da  $k = ch$  gilt, wenn  $c$  eine positive Konstante bezeichnet, ist der zweite Term auf der rechten Seite dieser Gleichung von der Größenordnung  $O(h)$ , und der Diffusionsterm des Lax–Friedrichs–Verfahrens ist bis auf Terme der Ordnung  $O(k^2, h^2)$  durch

$$\frac{h^2}{2k}u_{xx}$$

gegeben. Die Differenzbildung beider Verfahren, das heißt, Ausgangsdatensatz minus Ergebnis des Tiefpaßfilters, liefert daher den Diffusionsterm

$$-\frac{h^2}{2k}u_{xx}.$$

Im Diskreten ist dieser Ausdruck durch

$$-\frac{U_{i+1}^n - 2U_i^n + U_{i-1}^n}{2k}$$

gegeben. Der nichtlineare Filteralgorithmus, der im zweiten Schritt angewendet wird, besitzt die konservative Diskretisierung

$$\frac{U_i^{n+1} - U_i^n}{k} = \frac{1}{2h^2} ((d_{i+1} + d_i)(U_{i+1}^n - U_i^n) + (d_{i-1} + d_i)(U_{i-1}^n - U_i^n)),$$

wobei  $d_i = 1$  gewählt wird, wenn an der Stelle  $i$  ein Extremwert vorliegt, andernfalls wird  $d_i = 0$  gesetzt. Aus Gründen der Konsistenz wird der Korrekturterm, das heißt, die rechte Seite der Filtergleichung, bei der Implementierung zusätzlich mit  $h$  multipliziert, so daß der diskrete Diffusionsterm der Filtergleichung durch

$$\frac{1}{2h} ((d_{i+1} + d_i)(U_{i+1}^n - U_i^n) + (d_{i-1} + d_i)(U_{i-1}^n - U_i^n))$$

gegeben ist.

Das Zusammensetzen der einzeln berechneten Diffusionen liefert für die gesamte Filterprozedur mit vorheriger Datenzerlegung den Diffusionsterm

$$-\frac{U_{i+1}^n - 2U_i^n + U_{i-1}^n}{2k} + \frac{1}{2h} ((d_{i+1} + d_i)(U_{i+1}^n - U_i^n) + (d_{i-1} + d_i)(U_{i-1}^n - U_i^n)). \quad (4.34)$$

Bei der numerischen Implementierung ist zu beachten, daß diese Diffusion (4.34) noch zu dem tiefpaßgefilterten Datensatz, also zum Ergebnis des Lax–Friedrichs–Verfahrens, addiert

werden muß, um die gefilterte Version des Eingangsdatensatzes wiederzugeben.

Darüber hinaus bietet sich die Möglichkeit, die berechnete Diffusion (4.34) direkt in das Lax–Friedrichs–Verfahren zu implementieren, so daß sich das neue Differenzenverfahren

$$\begin{aligned}
 U_i^{n+1} &= U_i^n - \frac{k}{2h} (f(U_{i+1}^n) - f(U_{i-1}^n)) + \frac{1}{2} (U_{i+1}^n - 2U_i^n + U_{i-1}^n) \\
 &\quad - \frac{U_{i+1}^n - 2U_i^n + U_{i-1}^n}{2} + \frac{k}{2h} ((d_{i+1} + d_i)(U_{i+1}^n - U_i^n) + (d_{i-1} + d_i)(U_{i-1}^n - U_i^n)) \\
 &= U_i^n - \frac{k}{2h} (f(U_{i+1}^n) - f(U_{i-1}^n)) \\
 &\quad + \frac{k}{2h} ((d_{i+1} + d_i)(U_{i+1}^n - U_i^n) + (d_{i-1} + d_i)(U_{i-1}^n - U_i^n))
 \end{aligned} \tag{4.35}$$

ergibt.

Numerische Ergebnisse für diese Art der Implementierung werden in Abschnitt 4.2.4 dargestellt.

Bevor wir auf die Numerik eingehen, leiten wir zunächst die kontinuierliche Form des aus Gesamtdiffusion (4.34) und Differenzenschema zusammengesetzten Verfahrens (4.35) her. Dabei ist eine Fallunterscheidung für die verschiedenen Kombinationen der Vorfaktoren  $d_{i-1}$ ,  $d_i$  und  $d_{i+1}$  notwendig. Exemplarisch zeigen wir die ausführliche Herleitung der kontinuierlichen Gleichung für den Fall  $d_{i-1} = d_i = d_{i+1} = 1$ :

Das Zentrale–Differenzen–Verfahren mit direkt integrierter Diffusion aus der Datenzerlegung ist durch (4.35) gegeben, besteht also aus dem Lax–Friedrichs–Verfahren (als Ergebnis des Tiefpaßfilters) und der Diffusion aus der Datenzerlegung. Für  $d_{i-1} = d_i = d_{i+1} = 1$  läßt sich dieser Ausdruck zu

$$U_i^{n+1} = U_i^n - \frac{k}{2h} (f(U_{i+1}^n) - f(U_{i-1}^n)) + \frac{k}{h} (U_{i+1}^n - 2U_i^n + U_{i-1}^n)$$

vereinfachen, und weitere Termumformung liefert

$$\frac{U_i^{n+1} - U_i^n}{k} = -\frac{f(U_{i+1}^n) - f(U_{i-1}^n)}{2h} + \frac{h^2}{h} \frac{U_{i+1}^n - 2U_i^n + U_{i-1}^n}{h^2}.$$

Dieser Ausdruck läßt sich in kontinuierlicher Form bis auf Terme der Ordnung  $O(k, h^2)$  als

$$\begin{aligned}
 u_t &= -f(u)_x + h \cdot u_{xx} \\
 \Leftrightarrow u_t + f(u)_x &= hu_{xx}
 \end{aligned}$$

darstellen und ist damit eine Advektions–Diffusions–Gleichung vom Typ

$$u_t + Au_x = Du_{xx}$$

mit  $A = u$  und  $D = h$ .

Gleiches Vorgehen ist auch auf alle anderen Kombinationsmöglichkeiten der Vorfaktoren  $d_{i-1}$ ,  $d_i$ ,  $d_{i+1}$  anwendbar, so daß sich für das Differenzenverfahren mit integrierter Filterroutine bis auf Terme der Größenordnung  $O(h^2)$  und  $O(k)$  die folgenden kontinuierlichen

Formen ergeben:

$$\begin{aligned}
(1) \quad & d_{i-1} = d_i = d_{i+1} = 1 \quad : \quad u_t + f(u)_x = hu_{xx}, \\
(2) \quad & d_{i-1} = d_i = d_{i+1} = 0 \quad : \quad u_t + f(u)_x = 0, \\
(3) \quad & d_{i-1} = 1, d_i = d_{i+1} = 0 \quad : \quad u_t + f(u)_x = -\frac{1}{2}u_x + \frac{h}{4}u_{xx}, \\
(4) \quad & d_{i-1} = d_{i+1} = 0, d_i = 1 \quad : \quad u_t + f(u)_x = \frac{h}{2}u_{xx}, \\
(5) \quad & d_{i-1} = d_i = 0, d_{i+1} = 1 \quad : \quad u_t + f(u)_x = \frac{1}{2}u_x + \frac{h}{4}u_{xx}, \\
(6) \quad & d_{i-1} = 0, d_i = d_{i+1} = 1 \quad : \quad u_t + f(u)_x = \frac{1}{2}u_x + \frac{3h}{4}u_{xx}, \\
(7) \quad & d_{i-1} = d_{i+1} = 1, d_i = 0 \quad : \quad u_t + f(u)_x = \frac{h}{2}u_{xx}, \\
(8) \quad & d_{i-1} = d_i = 1, d_{i+1} = 0 \quad : \quad u_t + f(u)_x = -\frac{1}{2}u_x + \frac{3h}{4}u_{xx}.
\end{aligned} \tag{4.36}$$

Bemerkenswert bei dieser Auflistung der kontinuierlichen Gleichungen ist die Tatsache, daß die Ableitung  $u_x$  in den Fällen (3), (5), (6) und (8) ohne einen Faktor der Größenordnung  $h$  oder  $h^2$  auftritt, da diese Terme selbst bei Grenzbetrachtungen für  $h \rightarrow 0$  erhalten bleiben.

Für weiterführende mathematische Betrachtungen ist es sinnvoll, die kontinuierlichen Gleichungen bis zur Ordnung  $O(k^2)$  zu betrachten und nicht schon, wie bisher, Terme der Größenordnung  $O(k)$  zu vernachlässigen, da aufgrund der CFL-Bedingung zwischen  $h$  und  $k$  der lineare Zusammenhang  $k = c \cdot h$  mit konstantem  $c > 0$  besteht.

Mit Hilfe der Taylorentwicklung ergibt sich unter dieser Voraussetzung aus dem Lax–Friedrichs–Verfahren

$$\frac{U_i^{n+1} - U_i^n}{k} = -\frac{1}{2h} (f(U_{i+1}^n) - f(U_{i-1}^n)) + \frac{1}{2k} (U_{i+1}^n - 2U_i^n + U_{i-1}^n)$$

die Differentialapproximation

$$u_t + \frac{k}{2}u_{tt} + f(u)_x = \frac{h^2}{2k}u_{xx} + O(h^2) + O(k^2), \tag{4.37}$$

so daß die kontinuierliche Gleichung des Lax–Friedrichs–Verfahrens bis auf Terme der Größenordnung  $O(h^2)$  und  $O(k^2)$  durch

$$u_t + \frac{k}{2}u_{tt} + f(u)_x = \frac{h^2}{2k}u_{xx} \tag{4.38}$$

gegeben ist.

Diese Darstellung entspricht der in Shokin [58] eingeführten  $\Gamma$ -Form der Differentialdarstellung des zugehörigen Differenzenschemas. Shokin konstruiert sie aus dem Differenzenschema, indem er die diskreten Differenzen durch ihre entsprechenden Differentialapproximationen ersetzt. Handelt es sich bei dem Differenzenschema um ein Verfahren der Ordnung  $\gamma_1$  bezüglich  $t$  und  $\gamma_2$  bezüglich  $x$  oder der Ordnung  $\gamma$ , wenn, wie im vorliegenden Fall,  $h = h(k)$  angenommen wird, und werden in der  $\Gamma$ -Form die Terme der Ordnung  $O(k^{\gamma_1+1}, h^{\gamma_2+1})$  beziehungsweise

$O(h^{\gamma+1})$  vernachlässigt, ergibt sich die von Shokin betrachtete  $\Gamma$ -Form der ersten Differentialapproximation des gegebenen Differenzenschemas.

Da das Lax-Friedrichs-Verfahren ein Differenzenschema erster Ordnung ist, kann aus (4.37) gefolgert werden, daß Gleichung (4.38), also

$$u_t + \frac{k}{2}u_{tt} + f(u)_x - \frac{h^2}{2k}u_{xx} = 0,$$

gleichzeitig die  $\Gamma$ -Form der ersten Differentialapproximation des Lax-Friedrichs-Schemas darstellt.

Analog ergeben sich die  $\Gamma$ -Formen der ersten Differentialapproximationen des Zentrale-Differenzen-Verfahrens mit direkt integrierter Gesamtdiffusion aus der Datenzerlegung. Für die Fälle (1) bis (8) aus (4.36) ergibt sich jeweils auf der linken Gleichungsseite, ebenso wie im Lax-Friedrichs-Verfahren, der zusätzliche Summand

$$\frac{1}{2}ku_{tt}.$$

Aus der  $\Gamma$ -Form läßt sich die ebenfalls von Shokin definierte  $\Pi$ -Form herleiten. Allgemein ergibt sich die  $\Pi$ -Form eines Differenzenverfahrens aus der  $\Gamma$ -Form, indem die Ableitungen  $\frac{\partial^l u}{\partial t^l}$  für  $l \geq 2$  durch Ableitungen nach  $x$  mit Hilfe der  $\Gamma$ -Form ersetzt werden. Besitzt das ursprüngliche Differenzenverfahren die Approximationsordnung  $\gamma_1$  beziehungsweise  $\gamma_2$  bezüglich  $t$  beziehungsweise  $x$ , und vernachlässigt man in der  $\Pi$ -Form Terme der Ordnung  $O(k^{\gamma_1+1}, h^{\gamma_2+1})$ , so ergibt sich die  $\Pi$ -Form der ersten Differentialapproximation des zugehörigen Differenzenschemas. Entscheidend bei der Entwicklung der  $\Pi$ -Form eines Differenzenverfahrens ist, daß dabei nicht auf die zu approximierende Differentialgleichung zurückgegriffen wird. Da das Differenzenschema die exakte Differentialgleichung nicht mehr erfüllt, werden damit, im Gegensatz zu anderen mathematischen Entwicklungsmethoden, zusätzliche Fehlerquellen vermieden. Allerdings ist zu beachten, daß sich die  $\Pi$ -Form der ersten Differentialapproximation genauso durch das formale Ersetzen der Ableitungen nach  $t$  bis zur Ordnung  $\gamma_1$  durch Ableitungen nach  $x$  unter Verwendung der zugrundeliegenden Differentialgleichung ergibt. Unterschiede in der resultierenden  $\Pi$ -Form treten erst ab der zweiten Differentialapproximation auf.

Zur Entwicklung der  $\Pi$ -Form der ersten Differentialapproximation des Lax-Friedrichs-Verfahrens können also entweder verschiedene Differentialoperatoren so auf die  $\Gamma$ -Form angewendet werden, daß bei Addition der resultierenden Gleichungen die Zeitableitungen sowie die gemischten Ableitungen eliminiert werden, oder diese Ableitungen werden unter Verwendung der zugehörigen Differentialgleichung ersetzt. Letztere Möglichkeit wird aufgrund ihrer Kürze für die vorliegende Situation gewählt:

Aus der Differentialgleichung

$$u_t = -uu_x$$

ergibt sich

$$\begin{aligned} u_{tt} &= (-uu_x)_t = -u_t u_x - uu_{tx} = -(-uu_x)u_x - u(-uu_x)_x \\ &= uu_x^2 + u^2 u_{xx} + uu_x^2 = 2uu_x^2 + u^2 u_{xx}, \end{aligned}$$

und Einsetzen dieses Ausdrucks in die  $\Gamma$ -Form (4.38) liefert

$$u_t + f(u)_x = -kuu_x^2 + \left( \frac{h^2}{2k} - \frac{k}{2}u^2 \right) u_{xx}, \quad (4.39)$$

die  $\Pi$ -Form der ersten Differentialapproximation des Lax–Friedrichs–Verfahrens. Mit der Bezeichnung  $\tau = \frac{k}{h}$  läßt sich die  $\Pi$ -Form weiterhin zu der kompakteren und bei Shokin verwendeten Darstellung

$$u_t + f(u)_x = \frac{h^2}{2k} [(1 - \kappa^2 u^2) u_x]_x \quad (4.40)$$

umformen.

Analog zum Vorgehen für das Lax–Friedrichs–Verfahren berechnen sich die  $\Pi$ -Formen für die Fälle (1) bis (8) des Differenzenverfahrens mit integriertem Diffusionsterm. Stets wird im zusätzlichen Term  $\frac{1}{2}ku_{tt}$  der Ausdruck  $u_{tt}$  durch  $2uu_x^2 + u^2u_{xx}$  ersetzt, so daß in den Differentialapproximationen der Fälle (1) bis (8) jeweils auf der rechten Seite der Term

$$-kuu_x^2 - \frac{k}{2}u^2u_{xx}$$

hinzukommt. Die  $\Pi$ -Formen sind damit für die verschiedenen Fälle durch

$$\begin{aligned} (1) \, d_{i-1} = d_i = d_{i+1} = 1 & : \quad u_t + f(u)_x = -kuu_x^2 + (h - \frac{k}{2}u^2)u_{xx}, \\ (2) \, d_{i-1} = d_i = d_{i+1} = 0 & : \quad u_t + f(u)_x = -kuu_x^2 - \frac{k}{2}u^2u_{xx}, \\ (3) \, d_{i-1} = 1, \, d_i = d_{i+1} = 0 & : \quad u_t + f(u)_x = -(\frac{1}{2} + kuu_x)u_x + (\frac{h}{4} - \frac{k}{2}u^2)u_{xx}, \\ (4) \, d_{i-1} = d_{i+1} = 0, \, d_i = 1 & : \quad u_t + f(u)_x = -kuu_x^2 + (\frac{h}{2} - \frac{k}{2}u^2)u_{xx}, \\ (5) \, d_{i-1} = d_i = 0, \, d_{i+1} = 1 & : \quad u_t + f(u)_x = (\frac{1}{2} - kuu_x)u_x + (\frac{h}{4} - \frac{k}{2}u^2)u_{xx}, \\ (6) \, d_{i-1} = 0, \, d_i = d_{i+1} = 1 & : \quad u_t + f(u)_x = (\frac{1}{2} - kuu_x)u_x + (\frac{3h}{4} - \frac{k}{2}u^2)u_{xx}, \\ (7) \, d_{i-1} = d_{i+1} = 1, \, d_i = 0 & : \quad u_t + f(u)_x = -kuu_x^2 + (\frac{h}{2} - \frac{k}{2}u^2)u_{xx}, \\ (8) \, d_{i-1} = d_i = 1, \, d_{i+1} = 0 & : \quad u_t + f(u)_x = -(\frac{1}{2} + kuu_x)u_x + (\frac{3h}{4} - \frac{k}{2}u^2)u_{xx} \end{aligned} \quad (4.41)$$

gegeben. An den  $\Pi$ -Formen der Fälle (1), (2), (4) und (7) ist sofort die Konsistenz der durch Datenzerlegung gefilterten Gleichung mit der zugrundeliegenden Differentialgleichung zu erkennen, da die rechten Seiten der Gleichungen jeweils aus Termen der Ordnung  $O(h)$  bestehen, die für  $h \rightarrow 0$  gegen Null konvergieren. Für die andern Fälle ist diese Aussage nicht unbedingt gültig, da einerseits ihre rechten Seiten für  $h \rightarrow 0$  nicht gegen Null konvergieren, und sich andererseits das Zusammenspiel der verschiedenen möglichen Fälle nicht überblicken läßt, so daß wir nicht analysieren können, welche Differentialgleichung in der Summe approximiert wird.

Ebenso wenig lassen sich Stabilitätsaussagen aus den berechneten  $\Pi$ -Formen ableiten. Für die bekannte Wärmeleitungsgleichung  $u_t + f(u)_x = Du_{xx}$  mit  $D > 0$  würde Stabilität folgen, aber im vorliegenden Fall treten Terme, die aus ersten Ableitungen bestehen, auf, und die Vorfaktoren der zweiten Ableitungen sind nicht notwendig positiv.

Insgesamt haben wir mit Hilfe einer Diffusionsanalyse der einzelnen Teilschritte beim Filtern

mit vorheriger Datenzerlegung eine Gesamtdiffusion entwickelt, die den Filter- und Zerlegungsvorgang mathematisch widerspiegelt und deren direkte Implementierung eine weitere numerische Anwendungsvariante liefert. Darüber hinaus lassen sich die möglichen kontinuierlichen Formen der resultierenden diskreten Gleichungen entwickeln und in die  $\Gamma$ - und  $\Pi$ -Form umformulieren. Eine daraus erhoffte Analyse des neuen Filterverfahrens bleibt allerdings aus.

Bevor wir auf die numerischen Anwendungen der neu entwickelten Filtermethode eingehen, vergleichen wir im folgenden Abschnitt das neue Filterverfahren mit dem aus der Bildverarbeitung bekannten Soft Wavelet Shrinkage.

### 4.2.3 Zusammenhang zum Soft Wavelet Shrinkage

Das Soft Wavelet Shrinkage [15] ist eine in der Bildverarbeitung häufig genutzte Technik, die zum Entrauschen von Signalen und Bildern bei gleichzeitiger Erhaltung ihrer Unstetigkeiten eingesetzt wird. In neuesten Veröffentlichungen von Steidl, Weickert und Mitarbeitern [63, 65] wird die Äquivalenz dieser Technik mit anderen, in der Bildverarbeitung bekannten Methoden wie der Totalvariations-Regularisierung und der Totalvariations-Diffusion gezeigt. Auch die von uns entwickelte Filtertechnik mit vorheriger Datenzerlegung weist gewisse Ähnlichkeiten mit dem Soft Wavelet Shrinkage auf, die im Folgenden näher untersucht werden. Es wird sich zeigen, daß eine quantitative Äquivalenz nicht gegeben ist, die Verfahren aber qualitativ sehr ähnlich sind.

Der Einfachheit halber bleiben wir für die Äquivalenzuntersuchungen im Eindimensionalen und betrachten nur eine spezielle Art von Wavelets, die Haar-Wavelets. Zunächst erklären wir das Prinzip des Soft Haar Wavelet Shrinkage.

#### Soft Haar Wavelet Shrinkage

Wir betrachten im Folgenden einen einzelnen Haar Wavelet Shrinkage-Schritt. Sei dazu  $f = (f_i)_{i=0}^{N-1}$  ein Eingangssignal und gelte  $N = 2^n$ . Dieses Eingangssignal wird in zwei Teilsignale  $c_i$  und  $b_i$  zerlegt, wobei

$$c_i = \frac{f_{2i} + f_{2i+1}}{\sqrt{2}}, \quad b_i = \frac{f_{2i} - f_{2i+1}}{\sqrt{2}} \quad (i = 0, \dots, \frac{N}{2} - 1)$$

gilt. Die Zerlegung heißt Analyseschritt,  $c_i$  sind die Koeffizienten einer Skalierungsfunktion und  $b_i$  die Koeffizienten des Wavelets auf dem nächst größeren Gitter. Die Mittelung

$$c_i = \frac{f_{2i} + f_{2i+1}}{\sqrt{2}}$$

stellt dabei das Ergebnis eines einfachen Tiefpaßfilters dar, bei dem der Nenner aus Skalierungsgründen  $\sqrt{2}$  statt 2 ist, und

$$b_i = \frac{f_{2i} - f_{2i+1}}{\sqrt{2}}$$

gibt das Ergebnis eines einfachen Hochpaßfilters zurück. Auf den Analyseschritt folgt die Shrinkage-Operation  $S_\tau(b_i)$ , die auf das Signal  $b_i$  angewendet wird und in Abhängigkeit des

Schwellenparameters  $\tau$  die Koeffizienten  $b_i$  schrumpft oder abschneidet.  $S_\tau$  bezeichnet im Allgemeinen eine nichtlineare, von  $\tau$  abhängige Funktion, die im Falle des Soft Shrinkage durch

$$S_\tau(\eta) = \begin{cases} \eta - \tau \operatorname{sgn} \eta & , \quad |\eta| \geq \tau \\ 0 & , \quad |\eta| < \tau \end{cases}$$

gegeben ist. Sie bewirkt, daß  $\eta$  entweder auf Null gesetzt wird oder aber um den Wert  $\tau$  reduziert wird. Die neuen Koeffizienten, die bei Anwendung von  $S_\tau$  auf  $b_i$  entstehen, bezeichnen wir mit  $\tilde{b}_i$ . Der letzte Schritt des Soft Wavelet Shrinkage ist der Syntheseschritt, in dem die Koeffizienten  $c_i$  und  $\tilde{b}_i$  zum neuen Signal  $u$  zusammengesetzt werden. Die Zusammensetzung erfolgt gemäß

$$\begin{aligned} u_{2i} &= \frac{c_i + S_\tau(b_i)}{\sqrt{2}} = \frac{f_{2i} + f_{2i+1}}{2} - \frac{1}{\sqrt{2}} S_\tau \left( \frac{f_{2i+1} - f_{2i}}{\sqrt{2}} \right), \\ u_{2i+1} &= \frac{c_i - S_\tau(b_i)}{\sqrt{2}} = \frac{f_{2i} + f_{2i+1}}{2} + \frac{1}{\sqrt{2}} S_\tau \left( \frac{f_{2i+1} - f_{2i}}{\sqrt{2}} \right). \end{aligned}$$

Das neue Signal  $u$  hat kleinere Wavelet-Koeffizienten und stellt die entrauschte Version des Eingangssignals  $f$  dar.

### Parallelen zum Filtern mit Datenzerlegung

Beim Filtern mit vorheriger Datenzerlegung können wir unseren durch Oszillationen veräuschten Datensatz ebenfalls mit  $f = (f_i)_{i=0}^{N-1}$  bezeichnen, wobei  $N$  die Anzahl der Gitterpunkte, auf denen der Algorithmus angewendet wird, angibt. Die Rolle des Tiefpaßfilters übernimmt das Lax-Friedrichs-Verfahren, das das überglättete Teilsignal  $c = (c_i)_{i=0}^{N-1}$  zurückgibt. Die Differenz aus dem Eingangssignal  $f$  und  $c$  liefert das Signal  $b = (b_i)_{i=0}^{N-1}$ , das als Ergebnis eines Hochpaßfilters interpretiert werden kann. Auf das Signal  $b$  wird der nichtlineare Filter

$$u_t = ((1 - \operatorname{sgn}(|u_x|))u_x)_x$$

angewendet. Die Anwendung dieses Filters auf die Daten  $b$  läßt sich als Shrinkage-Funktion

$$S_\tau(\eta, \eta_x, \eta_{xx}) = \eta + \tau(g(|\eta_x|)\eta_x)_x$$

mit  $g(|\eta_x|) = 1 - \operatorname{sgn}(|\eta_x|)$  darstellen, wobei für den Parameter  $\tau$  in unserem Fall  $\tau = \Delta t$  gilt. An dieser Stelle wird ein deutlicher Unterschied zum Soft Wavelet Shrinkage sichtbar. Die Funktion  $S_\tau$  hängt hier nicht nur von  $\eta$  sondern auch von den Ableitungen  $\eta_x$  und  $\eta_{xx}$  ab, so daß die direkte Berechnung von  $S_\tau(b_i)$  nicht mehr in vollständiger Analogie zum Soft Wavelet Shrinkage vorgenommen werden kann. Qualitativ bewirkt die Funktion  $S_\tau(\eta, \eta_x, \eta_{xx})$  allerdings etwas sehr Ähnliches wie die eigentliche Shrinkage-Funktion  $S_\tau(\eta)$ , auch hier werden die Extremwerte geschrumpft und alle anderen Werte bleiben unverändert. Im Fall des Filterns mit vorheriger Datenzerlegung wird das Abschneiden der Extrema allerdings zusätzlich mit der Krümmung an den entsprechenden Punkten gewichtet, so daß hochfrequente Oszillationen stärker geschrumpft werden als niederfrequente. Das Verhältnis zwischen Regionengröße der Oszillationen und Schrumpfungsbetrag ist damit nichtlinear. Beim Soft Wavelet Shrinkage werden die Extremwerte dagegen antiproportional zu ihrer Ausdehnung geschrumpft.



Auch der Syntheseschritt des Soft Wavelet Shrinkage findet sich beim Filtern mit vorheriger Datenzerlegung wieder, die geschrumpften Koeffizienten  $\tilde{b}_i$  werden zu den Koeffizienten  $c_i$  aus dem Lax–Friedrichs–Verfahren addiert, und es ergibt sich der neue, gefilterte Datensatz  $u$ .

Insgesamt betrachtet, läßt sich das Filterprinzip mit vorheriger Datenzerlegung als Soft Haar Wavelet Shrinkage interpretieren, wobei wir allerdings einen anderen Tiefpaßfilter und eine andere Shrinkage–Funktion verwenden.

Der enge Zusammenhang zwischen unserer neu entwickelten Filtermethode und dem aus der Bildverarbeitung bekannten und häufig verwendeten Soft Wavelet Shrinkage unterstützt unsere Annahme, ein geeignetes Filterprinzip für das nichtlineare Glätten von Oszillationen in Lösungen hyperbolischer Erhaltungsgleichungen gefunden zu haben.

Die numerischen Anwendungen im folgenden Abschnitt belegen die gute Funktionalität des neu entwickelten Filterprinzips.

#### 4.2.4 Resultierende Algorithmen und ihre Anwendung

Zunächst implementieren wir das Filterprinzip mit vorheriger Datenzerlegung, ohne die Kenntnis der dabei entstehenden Gesamtdiffusion auszunutzen. Als erstes untersuchen wir dabei die Filterimplementierung nach Beendigung des Differenzenverfahrens.

##### Datenzerlegung nach Beendigung des Differenzenverfahrens

Wir lösen das eindimensionale Cauchy–Problem (1.49) für Burgers’ Gleichung mit dem Lax–Wendroff–Verfahren. Da es sich um ein Verfahren höherer Ordnung handelt, treten Oszillationen in der Umgebung der Unstetigkeit auf, die wir mit Hilfe der Datenzerlegung glätten wollen. Dazu berechnen wir parallel zur Lax–Wendroff–Lösung das entsprechende Ergebnis des Lax–Friedrichs–Verfahrens, das ein Differenzenverfahren erster Ordnung ist und damit als Tiefpaßfilter eingesetzt werden kann. Auf die Differenz aus Lax–Wendroff– und Lax–Friedrichs–Lösung wenden wir die nichtlineare Filterroutine

$$u_t = ((1 - \operatorname{sgn} |u_x|)u_x)_x \quad (4.42)$$

an und addieren anschließend den gefilterten Datensatz zu der Lax–Friedrichs–Lösung. Auf einem Rechengitter aus 400 äquidistanten Gitterpunkten ergibt sich nach  $t = 2.5$  bei einer CFL–Zahl von 0.5 und 20 Iterationen des nichtlinearen Filters (4.42) das in Abbildung 4.42 dargestellte Ergebnis. Zum Vergleich sind die zugehörigen Lösungen des Lax–Wendroff– und Lax–Friedrichs–Verfahrens ebenfalls dargestellt.

Die mit Hilfe der Datenzerlegung gefilterte Lösung zeigt, daß die Oszillationen des Lax–Wendroff–Verfahrens reduziert sind, ohne daß Überglättungen aufgrund der Anwendung des Lax–Friedrichs–Verfahrens zu erkennen sind. Die tatsächliche Höhe und die Lage der Unstetigkeit sind erhalten.

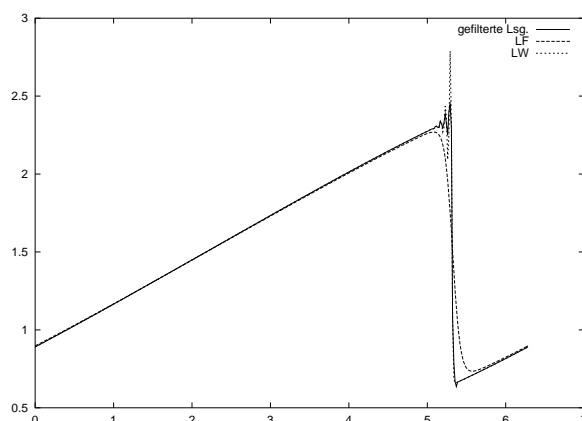


Abbildung 4.42: Mittels Datenzerlegung gefiltertes Ergebnis sowie Lax–Wendroff– und Lax–Friedrichs–Lösung

### Datenzerlegung nach jedem Zeitschritt des Differenzenverfahrens

Von größerer Bedeutung ist die Filterimplementierung nach jedem Zeitschritt des Differenzenverfahrens, der wir uns im Folgenden widmen.

Während jedes Zeitschrittes des Differenzenschemas wird dazu sowohl ein Schritt des zugrundeliegenden Lax–Wendroff–Verfahrens als auch ein Schritt des als Tiefpaß aufgefaßten Lax–Friedrichs–Verfahrens gelöst. Auf den Differenzdatensatz wird eine Filteriteration des nichtlinearen Filters angewendet, und das Ergebnis, addiert zu dem Datensatz aus dem Lax–Friedrichs–Schritt, liefert die Ausgangsdaten für den darauffolgenden Zeitschritt. Wenden wir dieses Verfahren wieder auf das einfache Cauchy–Problem (1.49) für Burgers’ Gleichung an, ergibt sich auf einem Rechengitter aus 400 äquidistanten Gitterpunkten bei einer CFL–Zahl von 0.6 nach  $t = 2.5$  die in Abbildung 4.43 links dargestellte Lösung. Die Oszillationen

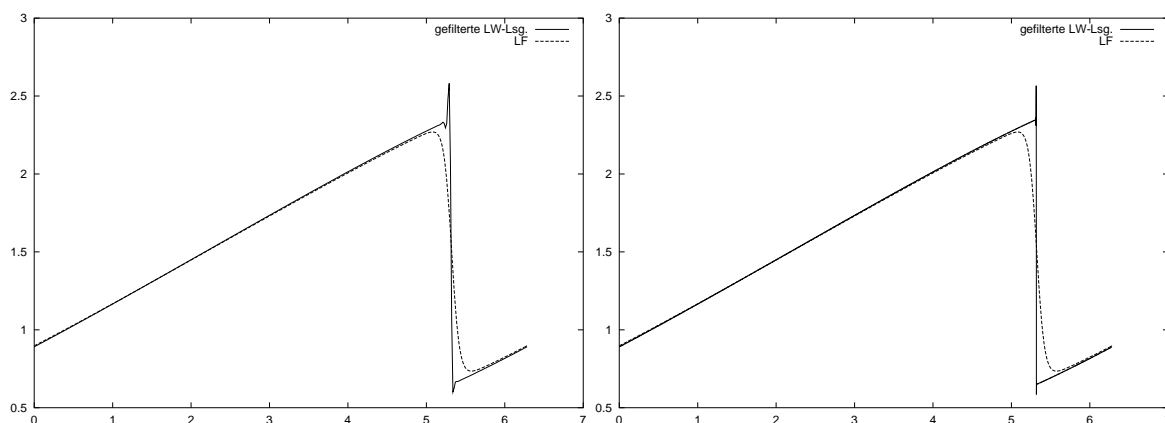


Abbildung 4.43: Nach jedem Zeitschritt gefilterte Lax–Wendroff–Lösung und zugehörige Lax–Friedrichs–Lösung für  $N = 400$  (links) und  $N = 6400$  (rechts)

des Lax–Wendroff–Verfahrens sind geglättet, nur direkt vor und hinter dem Stoß sind noch Restoszillationen zu erkennen. Sie können durch Anwendung mehrerer Filteriterationen pro Zeitschritt weiter verringert werden, sind aber nie vollständig geglättet. Ein Vergleich mit

der zugehörigen Lax–Friedrichs–Lösung zeigt, daß nicht nur die Höhe sondern auch die Lage der Unstetigkeit korrekt erhalten ist. Für  $h \rightarrow 0$  konvergiert das Verfahren, wie der rechte Teil der Abbildung für  $N = 6400$  andeutet.

Die Leistungsfähigkeit der neuen Filtermethode wird deutlich, wenn wir das stabile Lax–Wendroff–Verfahren durch das instabile Zentrale–Differenzen–Schema zweiter Ordnung ersetzen. Wie bereits bekannt, liefert dieses Verfahren bei einer CFL–Zahl von 0.009 nach  $t = 2.5$  eine sehr stark oszillierende Lösung, dargestellt in Abbildung 3.9, die schon wenige Zeitschritte später zusammenbricht. Abbildung 4.44 links zeigt dagegen die nach jedem Zeitschritt durch Datenzerlegung gefilterte Lösung zum gleichen Zeitpunkt, wobei eine wesentlich größere CFL–Zahl von 0.1 gewählt wurde. In der Lösung sind nur noch wenige Restoszillationen zu erkennen, und Höhe und Lage der Unstetigkeit sind korrekt erhalten, wie der Vergleich mit der zugehörigen Lax–Friedrichs–Lösung zeigt.

Betrachten wir allerdings das Lösungsverhalten für  $h \rightarrow 0$ , stellen wir fest, daß sich für  $N = 1600$ , also viermal so kleine Gitterweite, neue Oszillationen bilden, wie die rechte Darstellung in Abbildung 4.44 zeigt. Da dieses Verhalten beim Lax–Wendroff–Verfahren als zu-

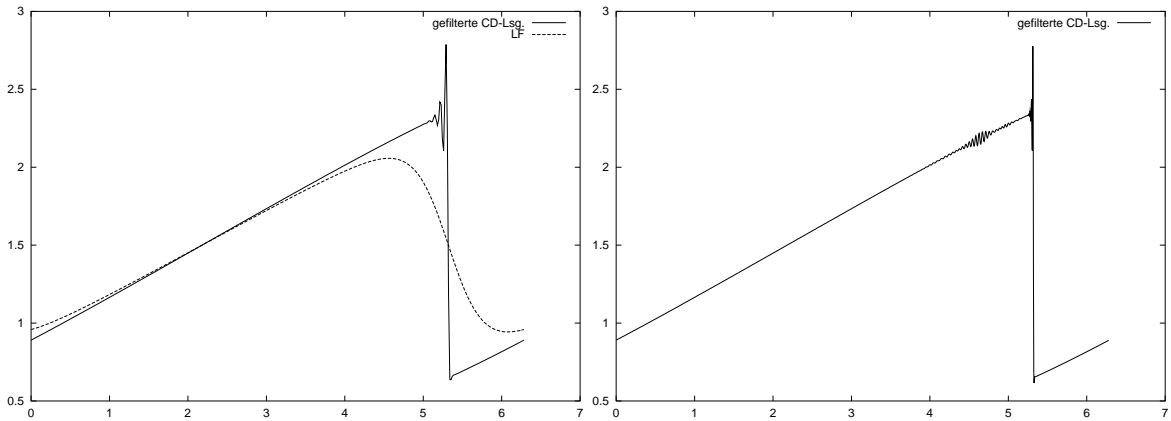


Abbildung 4.44: Nach jedem Zeitschritt gefilterte Zentrale–Differenzen–Lösung im Vergleich zur Lax–Friedrichs–Lösung für  $N = 400$  (links) und  $N = 1600$  (rechts)

grundlegendes Differenzenverfahren nicht auftritt, sind diese Oszillationen, ebenso wie in Kapitel 4.1.5, auf das instabile Zentrale–Differenzen–Verfahren zurückzuführen.

Das Entstehen der Oszillationen kann auch hier verhindert werden, indem wir eine kleinere CFL–Zahl, beispielsweise 0.05, wählen oder die Zeitschrittweite verkleinern. Ersetzen wir die hyperbolische durch eine parabolische Zeitschrittweite, bilden sich zwar keine Oszillationen mehr aus, aber das resultierende Verfahren benötigt sehr viel Rechenzeit. Reduzieren wir die hyperbolische Zeitschrittweite dagegen gemäß  $k^* = \frac{k}{10}$ , ist das Verfahren für  $h \rightarrow 0$  ebenfalls stabil und liefert die numerische Lösung mit weitaus geringerem Rechenaufwand. Abbildung 4.45 zeigt die so berechnete Lösung für  $N = 3200$ .

Zusammenfassend zeigen die numerischen Experimente, daß die Filteranwendung mittels Datenzerlegung für das Lax–Wendroff–Verfahren zu der gewünschten Reduktion der Oszillationen führt, wobei tatsächliche Unstetigkeiten in ihrer Höhe und Lage erhalten bleiben. Für das Zentrale–Differenzen–Verfahren zweiter Ordnung lassen sich diese Ziele ebenfalls erreichen, allerdings ist die Stabilität des resultierenden Verfahrens für  $h \rightarrow 0$  nur gegeben, wenn die Zeitschrittweite oder die CFL–Zahl hinreichend klein gewählt werden.

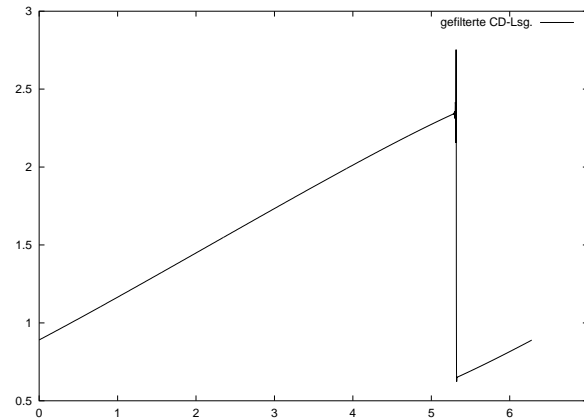


Abbildung 4.45: Nach jedem Zeitschritt gefilterte Zentrale-Differenzen-Lösung mit reduzierter Zeitschrittweite  $k^* = k/10$  für  $N = 3200$

### Direkte Implementierung der Gesamtdiffusion

Die Berechnung der Diffusion (4.34), die aus der gesamten Filterprozedur mittels Datenzerlegung resultiert, liefert die Möglichkeit, den entwickelten Diffusionsterm direkt in das zu filternde Differenzenverfahren zu implementieren. Dabei ist zu beachten, daß die berechnete Diffusion (4.34), in der bereits das zu filternde Zentrale-Differenzen-Verfahren enthalten ist, bei der direkten Implementierung in das Lax-Friedrichs-Schema integriert wird, da die Lax-Friedrichs-Lösung den tiefpaßgefilterten Teil der Lösung darstellt. Das resultierende Verfahren (4.35) läßt sich daher auf zwei Arten interpretieren, einerseits als von Oszillationen befreites Zentrale-Differenzen-Verfahren, andererseits als kantenverschärfendes Lax-Friedrichs-Verfahren.

### Eindimensionale skalare Testfälle

Wir beginnen die numerischen Anwendungen wieder mit dem einfachen Cauchy-Problem (1.49) für Burgers' Gleichung, das wir auf einem Gitter aus 400 äquidistanten Punkten bei einer CFL-Zahl von 0.1 mit (4.35) lösen. Nach der Zeit  $t = 2.5$  ergibt sich das in Abbildung 4.46 links zusammen mit der zugehörigen Lax-Friedrichs-Lösung dargestellte Ergebnis. Die Oszillationen des Zentrale-Differenzen-Verfahrens sind weitgehend geglättet, die Steilheit des Stoßes ist im Gegensatz zum Lax-Friedrichs-Verfahren erhalten und auch die Stoßlage stimmt mit den physikalischen Gegebenheiten überein. Auffällig sind allerdings die kleinen Oszillationen im linearen Anteil der Lösung, die bei  $h \rightarrow 0$  weiter anwachsen, was sich schon in Abbildung 4.46 rechts bei halb so großer Raumschrittweite andeutet. Das Entstehen der Oszillationen ist, wie schon bei der Datenzerlegung nach jedem Zeitschritt, zu vermeiden, indem wir die hyperbolische Zeitschrittweite durch eine parabolische ersetzen oder die hyperbolische Zeitschrittweite durch einen künstlich eingeführten Faktor verkleinern. Abbildung 4.47 links zeigt das Ergebnis der numerischen Simulation für eine gemäß  $k^* = \frac{k}{10}$  verkleinerte hyperbolische Zeitschrittweite. Die Oszillationen im linearen Lösungsanteil sind minimal, für  $h \rightarrow 0$  verschwinden sie ganz, und das Verfahren konvergiert, wie der rechte Teil der Abbildung für  $N = 1600$ , der vierfachen Anzahl von Gitterpunkten, andeutet.

Zurückzuführen sind die Oszillationen auf das instabile Zentrale-Differenzen-Verfahren, das

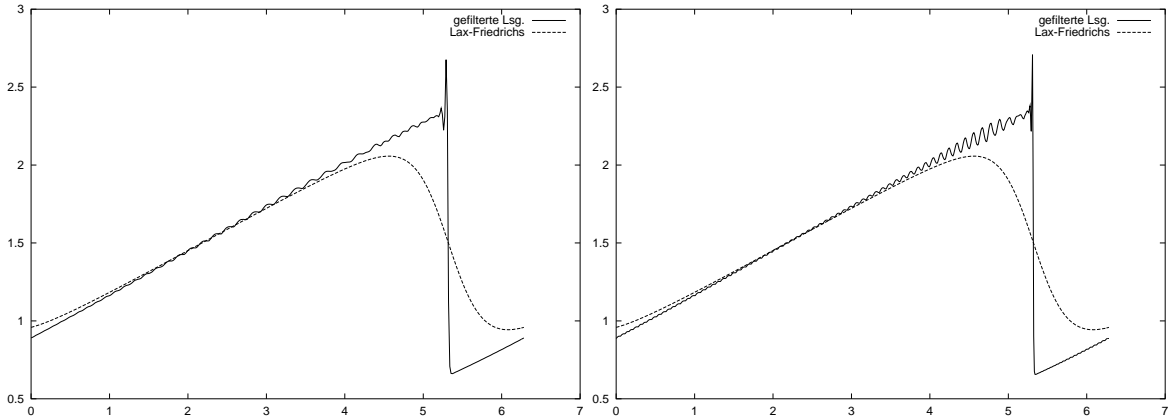


Abbildung 4.46: Zentrale-Differenzen-Lösung mit direkt implementierter Diffusion im Vergleich zur Lax-Friedrichs-Lösung bei  $N = 400$  (links) bzw.  $N = 800$  Gitterpunkten

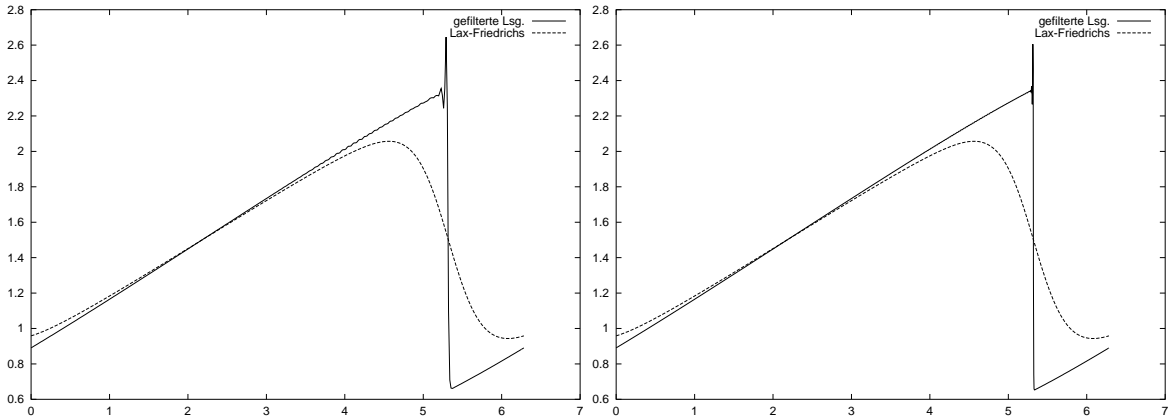


Abbildung 4.47: Zentrale-Differenzen-Lösung mit direkt implementierter Diffusion und reduzierter Zeitschrittweite bei  $N = 400$  (links) bzw.  $N = 1600$  Gitterpunkten

bei der Berechnung der Gesamtdiffusion (4.34) zugrundegelegt wurde, denn analoge Simulationen mit dem Lax-Wendroff-Verfahren als grundlegendes Differenzenschema erzeugen diese Oszillationen nicht, wie wir im Folgenden zeigen.

Ersetzen wir das Zentrale-Differenzen-Verfahren durch das Lax-Wendroff-Verfahren, ergibt sich bei der Differenz aus Lax-Wendroff- und Lax-Friedrichs-Schema bis auf Terme der Größenordnung  $O(k^2, h^2)$  der kontinuierliche Ausdruck

$$k u u_x^2 + \frac{k}{2} u^2 u_{xx} - \frac{h^2}{2k} u_{xx},$$

so daß wir die diskrete Gesamtdiffusion

$$\begin{aligned} & k U_i^n \left( \frac{U_i^n - U_{i-1}^n}{h} \right)^2 + \left( \frac{k}{2h^2} (U_i^n)^2 - \frac{1}{2k} \right) (U_{i+1}^n - 2U_i^n + U_{i-1}^n) \\ & + \frac{1}{2h} ((d_{i+1} + d_i)(U_{i+1}^n - U_i^n) + (d_{i-1} + d_i)(U_{i-1}^n - U_i^n)) \end{aligned} \quad (4.43)$$

erhalten. Integriert in das Lax–Friedrichs–Verfahren ergibt sich das Differenzenschema

$$\begin{aligned} U_i^{n+1} = & U_i^n - \frac{k}{2h} (f(U_{i+1}^n) - f(U_{i-1}^n)) \\ & + k^2 U_i^n \left( \frac{U_i^n - U_{i-1}^n}{h} \right)^2 + \frac{k^2}{2h^2} (U_i^n)^2 (U_{i+1}^n - 2U_i^n + U_{i-1}^n) \\ & + \frac{k}{2h} ((d_{i+1} + d_i)(U_{i+1}^n - U_i^n) + (d_{i-1} + d_i)(U_{i-1}^n - U_i^n)). \end{aligned}$$

Lösen wir damit das Cauchy–Problem (1.49) auf einem Gitter aus 400 äquidistanten Gitterpunkten bei einer CFL–Zahl von 0.5 erzielen wir das links in Abbildung 4.48 dargestellte Ergebnis. Es sind keine Oszillationen im linearen Teil der Lösung zu erkennen, und auch für kleinere Gitterweiten bilden sich keine neuen Oszillationen aus. Für  $h \rightarrow 0$  konvergiert das

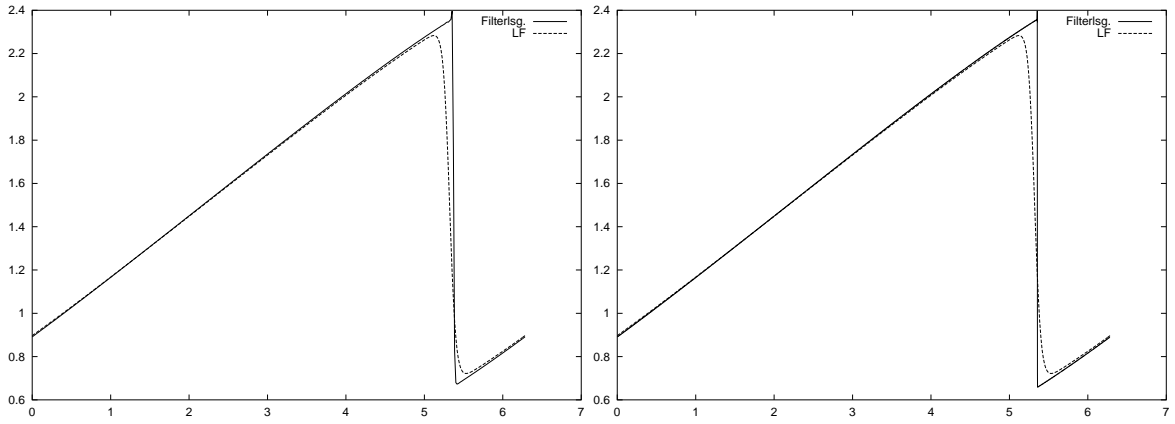


Abbildung 4.48: Lax–Wendroff–Lösung mit direkt implementierter Diffusion bei  $N = 400$  (links) bzw.  $N = 6400$  Gitterpunkten

Verfahren, wie der rechte Teil der Abbildung 4.48 für  $N = 6400$  andeutet. Die Qualität der Lösung ist schon für  $N = 400$  sehr beeindruckend, die Oszillationen sind bis auf einen sehr kleinen Überschwinger direkt vor dem Stoß geglättet, und die Unstetigkeit ist vollständig erhalten.

Darüber hinaus haben wir gezeigt, daß bei der Verkleinerung der Gitterweite keine neuen Oszillationen entstehen, so daß die Oszillationen, die bei der Implementierung von (4.35) für  $h \rightarrow 0$  auftreten, auf das instabile Zentrale–Differenzen–Verfahren zurückzuführen sind.

Bemerkenswert ist, daß selbst die direkte Implementierung des Diffusionstermes (4.34) ohne die Filterkoeffizienten, das heißt mit  $d_{i-1} = d_i = d_{i+1} = 1$ , also

$$\begin{aligned} & -\frac{U_{i+1}^n - 2U_i^n + U_{i-1}^n}{2k} + \frac{1}{h} (U_{i+1}^n - U_i^n + U_{i-1}^n - U_i^n) \\ = & \left( -\frac{1}{2k} + \frac{1}{h} \right) (U_{i+1}^n - 2U_i^n + U_{i-1}^n), \end{aligned}$$

eine gute Approximation der Lösung liefert, wie in Abbildung 4.49 für 400 Gitterpunkte und einer CFL–Zahl von 0.1 zu erkennen ist. Die Oszillationen sind vollständig geglättet und die Lage und Höhe der Unstetigkeit sind erhalten. Darüber hinaus treten bei kleiner werdender

Gitterweite keine Konvergenzprobleme bei der hyperbolischen Wahl der Zeitschrittweite  $k$  auf.

Zu erklären ist der gute Filtereffekt der Diffusion mit  $d_{i-1} = d_i = d_{i+1} = 1$  eventuell dadurch, daß schon der Ausdruck  $U_{i+1}^n - 2U_i^n + U_{i-1}^n$ , aus dem der Filterterm in diesem Fall hauptsächlich besteht, eine gewisse Filtersteuerung bewirkt. Liegen die Datenpunkte beispielsweise gleichmäßig verteilt auf einer Geraden, so gibt der Term den Wert Null zurück, und es wird nicht gefiltert, was in dieser Situation zu erwarten ist.

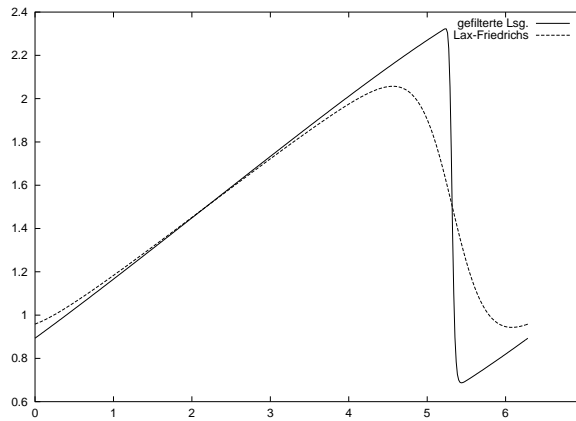


Abbildung 4.49: Mit  $d_{i-1} = d_i = d_{i+1} = 1$  gefilterte Zentrale-Differenzen-Lösung

### Eindimensionale Systeme von Erhaltungsgleichungen

Als weiteres Anwendungsbeispiel betrachten wir im Folgenden die direkte Implementierung der Gesamtdiffusion in Lösungsverfahren für die eindimensionalen Eulergleichungen und prüfen die Funktionalität dieses Filterprinzips für Systeme von Erhaltungsgleichungen.

Wir betrachten das einfache Riemann-Problem (1.58) ohne sonische Punkte, das eine Unstetigkeit an der Stelle  $x_0 = 0.3$  besitzt. Da es sich jetzt um ein System von Erhaltungsgleichungen handelt, integrieren wir die Gesamtdiffusion in jede Komponente des Gleichungssystems einzeln und unabhängig voneinander.

Implementieren wir für jeden Zeitschritt die errechnete Gesamtdiffusion (4.34), die auf dem Zentrale-Differenzen-Verfahren beruht, direkt in den tiefpaßgefilterten Datensatz, der durch das Lax-Friedrichs-Verfahren gegeben ist, ergeben sich gleiche Probleme wie im skalaren Fall. Auf einem Gitter aus 400 äquidistanten Punkten ergibt sich bei einer CFL-Zahl von 0.5 nach  $t = 0.2$  eine noch richtig erscheinende Dichte, die in Abbildung 4.50 links zusammen mit der entsprechenden Lax-Friedrichs-Lösung dargestellt ist. Für kleiner werdende Gitterweite entstehen aber neue Oszillationen, wie im rechten Teil von Abbildung 4.50 für  $N = 1600$  zu erkennen ist. Abhilfe schafft auch hier die künstliche Verkleinerung der hyperbolischen Zeitschrittweite gemäß  $k^* = \frac{k}{10}$ , wie die Dichteverläufe für  $N = 400$  und  $N = 1600$  in Abbildung 4.51 zeigen. Die Ergebnisse beeindrucken in ihrer Qualität, allerdings ist der Rechenaufwand durch die Verkleinerung der Zeitschrittweite erhöht.

In Tabelle 4.2 sind die numerischen Fehler und die experimentelle Konvergenzordnung des durch direkte Implementierung der Gesamtdiffusion (4.34) konstruierten Verfahrens mit reduzierter Zeitschrittweite für das Riemann-Problem (1.58) dargestellt. Die rechte Spalte zeigt die entsprechenden Werte, wenn an den Unstetigkeitsstellen ein Intervall der Länge

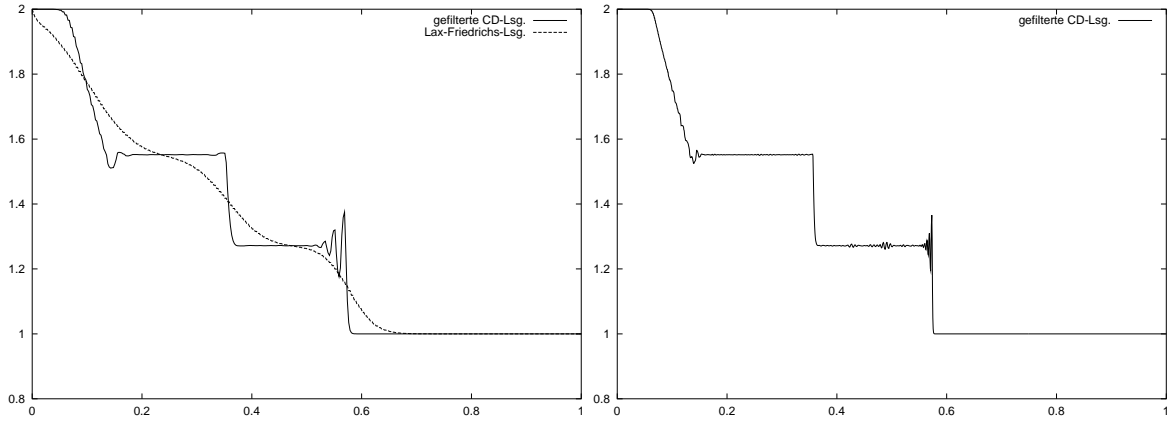


Abbildung 4.50: Mit direkt implementiertem Diffusionsterm gefilterte Zentrale-Differenzen-Lösung des Riemann-Problems (1.58) für  $N = 400$  (links) bzw.  $N = 1600$

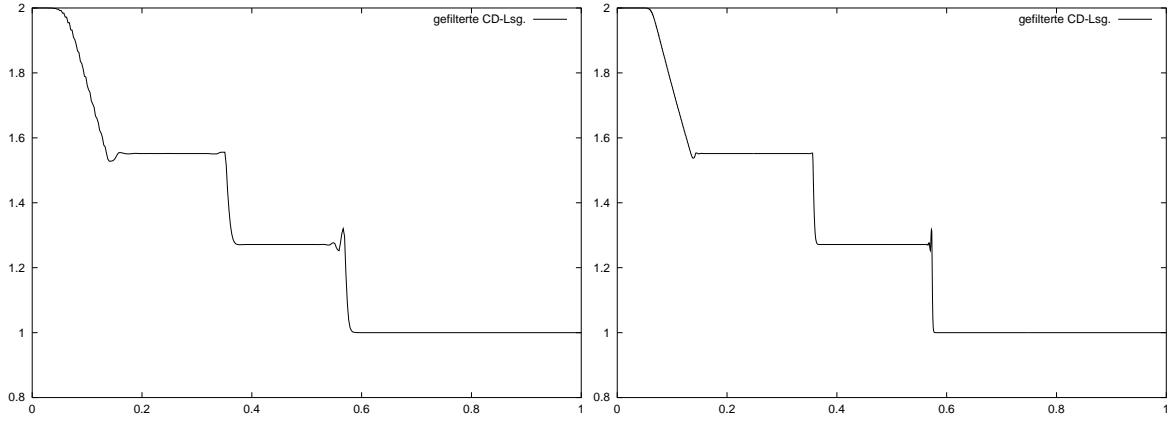


Abbildung 4.51: Mit direkt implementiertem Diffusionsterm gefilterte Zentrale-Differenzen-Lösung des Riemann-Problems (1.58) für  $N = 400$  (links) bzw.  $N = 1600$  bei reduzierter Zeitschrittweite

$\Delta I = 0.023$  bei der Berechnung unberücksichtigt bleibt. Da die EOC-Werte bei Ausschluß der Unstetigkeiten stets größer eins sind, ist das konstruierte Differenzenschema ein Verfahren zweiter Ordnung.

Selbst für das Riemann-Problem (1.59) mit sonischem Punkt liefert die direkte Implementierung der Gesamtdiffusion (4.34) mit reduzierter Zeitschrittweite ein stabiles Verfahren, das auf einem Gitter aus 400 Punkten bei der CFL-Zahl 0.5 nach  $t = 0.2$  die in Abbildung 4.52 dargestellte Dichte liefert. Die numerische Lösung weist zwar noch Restoszillationen auf, approximiert aber die Unstetigkeiten und sogar die Umgebung des sonischen Punktes ohne zusätzlichen Entropie-Fix sehr gut. Besonders vor dem Hintergrund, daß das Lax-Wendroff-Verfahren für dieses Testproblem instabil ist, beeindruckt die Leistungsfähigkeit des neu konstruierten Verfahrens, das eine gefilterte Version des Zentrale-Differenzen-Verfahrens darstellt.

Wir können das Reduzieren der Zeitschrittweite umgehen, wenn wir die Gesamtdiffusion



h	Verfahren (4.35)		Verfahren (4.35) ohne Unstetigk.	
	$\ u - u_h\ _{L_1}$	EOC	$\ u - u_h\ _{L_1}$	EOC
0.02	0.026110		0.018687	
0.01	0.012715	1.04	0.008685	1.04
0.005	0.007364	0.79	0.004117	1.08
0.0025	0.003949	0.90	0.001811	1.19
0.00125	0.001905	1.05	0.000774	1.23

Tabelle 4.2: Experimentelle Konvergenzordnung des durch Implementierung der Gesamtdiffusion (4.34) konstruierten Verfahrens bei reduzierter Zeitschrittweite

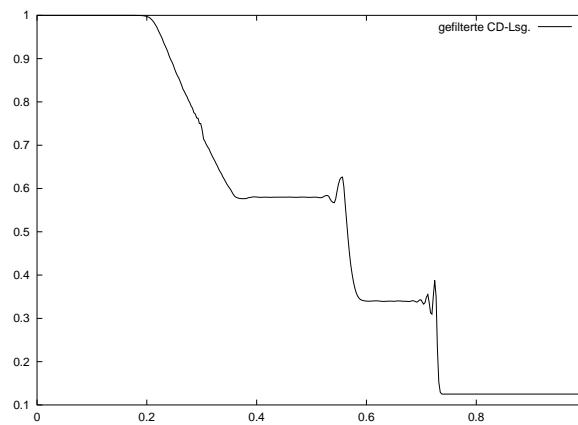


Abbildung 4.52: Mit direkt implementiertem Diffusionsterm gefilterte Zentrale-Differenzen-Lösung des Riemann-Problems (1.59) für  $N = 400$  bei reduzierter Zeitschrittweite

(4.43), die nicht mehr auf dem Zentrale-Differenzen- sondern auf dem Lax-Wendroff-Verfahren beruht, implementieren.

Die analogen Simulationen wie für (4.34) liefern für das Riemann-Problem (1.58) für  $N = 400$  und  $N = 1600$  die in Abbildung 4.53 dargestellten Dichteverläufe. Die Oszillationen sind vollständig geglättet, die Unstetigkeiten sind in Höhe und Lage erhalten, und das Verfahren konvergiert für  $h \rightarrow 0$ . Lassen wir den Term

$$kU_i^n(U_i^n - U_{i-1}^n)^2/h^2 \quad (4.44)$$

bei der Implementierung der Gesamtdiffusion (4.43) weg, ergeben sich noch glattere Ergebnisse, wie Abbildung 4.54 für  $N = 400$  und  $N = 1600$  Gitterpunkte verdeutlicht. Für  $N = 1600$  ist die approximative Lösung zusammen mit der exakten Lösung dargestellt. Tabelle 4.3 zeigt jedoch, daß die experimentelle Konvergenzordnung bei der Implementierung der Gesamtdiffusion (4.43), auch beim Weglassen des Terms (4.44), weitaus kleiner ist als bei den bisherigen Verfahren. Weitere Simulationen zeigen, daß die EOC-Werte selbst beim Aussparen der Unstetigkeitsstellen kleiner als eins bleiben, so daß das neue Differenzenverfahren lediglich von erster Ordnung genau ist.

Bei der Implementierung der Gesamtdiffusion (4.43) für das Riemann-Problem (1.59) mit sonischem Punkt ergibt sich kein stabiles Verfahren.

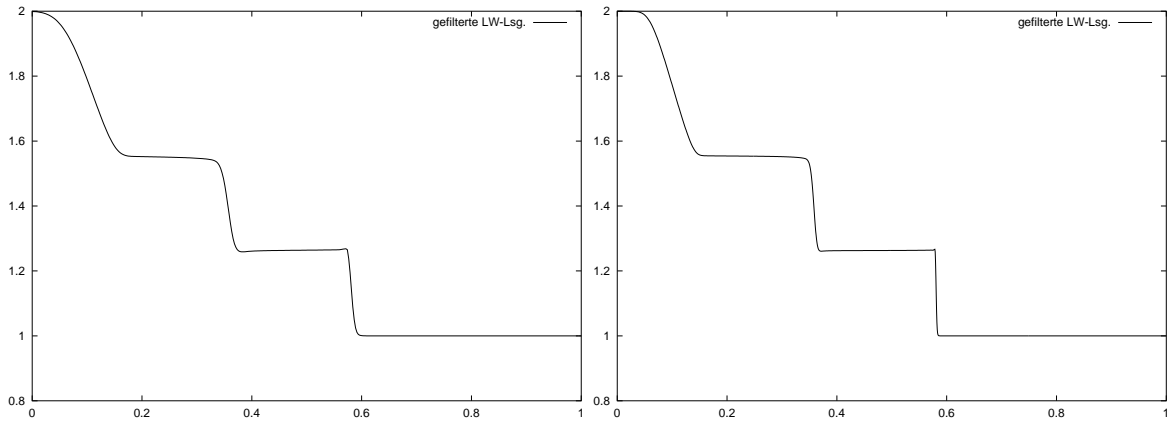


Abbildung 4.53: Mit direkt implementiertem Diffusionsterm (4.43) gefilterte Lax–Wendroff–Lösung des Riemann–Problems (1.58) für  $N = 400$  (links) bzw.  $N = 1600$

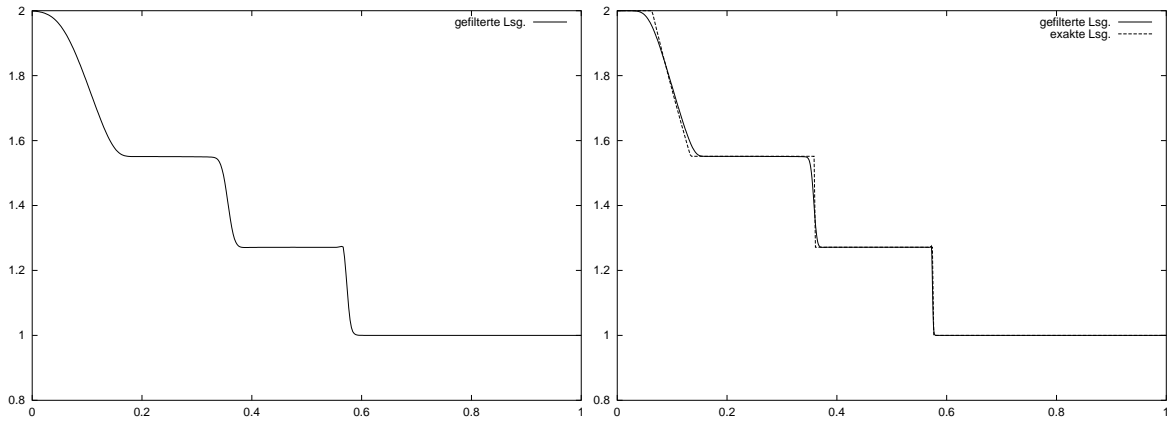


Abbildung 4.54: Mit direkt implementierter Diffusion (4.43) ohne (4.44) gefilterte Lax–Wendroff–Lösung des Riemann–Problems (1.58) für  $N = 400$  (links) bzw.  $N = 1600$

h	Impl. von (4.43)		Impl. von (4.43) ohne (4.44)	
	$\ u - u_h\ _{L_1}$	EOC	$\ u - u_h\ _{L_1}$	EOC
0.02	0.033481		0.031016	
0.01	0.024156	0.47	0.020684	0.58
0.005	0.016525	0.55	0.013640	0.60
0.0025	0.011882	0.48	0.008710	0.65
0.00125	0.009237	0.36	0.005474	0.67

Tabelle 4.3: Experimentelle Konvergenzordnung des durch Implementierung der Gesamtdiffusion (4.43) konstruierten Verfahrens mit und ohne dem Term (4.44)

### Zusammenfassung der numerischen Ergebnisse

Die numerischen Simulationen zeigen, daß das Filtern mit vorheriger Datenzerlegung für das numerische Lösen hyperbolischer Erhaltungsgleichungen anwendbar ist. Sowohl nach Be-

endigung eines Differenzenverfahrens höherer Ordnung als auch nach jedem Zeitschritt des Differenzenschemas bewirkt das Filterprinzip eine Glättung der Oszillationen bei weitgehender Erhaltung der Stoßhöhe, und die Lage des Stoßes ist stets korrekt dargestellt.

Die Filteranwendung nach Beendigung des Differenzenverfahrens liefert qualitativ nicht so gute Ergebnisse, da die Oszillationen nicht stark genug geglättet werden.

Bei der Filteranwendung mit vorheriger Datenzerlegung nach jedem Zeitschritt des Lax–Wendroff–Verfahrens entsteht ein stabiles Verfahren, das, bis auf eine Restoszillation direkt am Stoß, geglättete Lösungen liefert und für  $h \rightarrow 0$  konvergiert. Ersetzen wir das Lax–Wendroff–Verfahren durch das Zentrale–Differenzen–Schema zweiter Ordnung, erhalten wir zunächst sehr beeindruckende Ergebnisse, stellen bei Verkleinerung der Gitterweite allerdings fest, daß neue Oszillationen entstehen. Durch eine reduzierte Zeitschrittweite lassen sich diese Instabilitäten verhindern, so daß damit selbst das gefilterte Zentrale–Differenzen–Verfahren geglättete und stabile numerische Lösungen erzeugt. Allerdings ist die Rechengeschwindigkeit des resultierenden Verfahrens durch die Reduktion der Zeitschrittweite etwas herabgesetzt.

Gleiches gilt für die direkte Implementierung der berechneten Gesamtdiffusion (4.34). Um neue Oszillationen zu vermeiden, ist ebenfalls eine künstliche Verkleinerung der Zeitschrittweite notwendig, mit der das resultierende Verfahren, das als gefiltertes Zentrale–Differenzen–Schema aufzufassen ist, beachtliche numerische Ergebnisse erzielt. Selbst für Stoßrohrprobleme mit und ohne sonischem Punkt ergeben sich stabile Lösungen mit nur geringen Restoszillationen, in denen die Unstetigkeiten in Höhe und Lage erhalten sind. Die Berechnung des EOC–Wertes zeigt, daß es sich um ein Verfahren zweiter Ordnung handelt.

Um die Reduktion der Zeitschrittweite zu umgehen, haben wir die entsprechende Gesamtdiffusion (4.43), die auf dem Lax–Wendroff–Verfahren beruht, berechnet und direkt in das Differenzenverfahren implementiert. Das resultierende Verfahren ist stabil und konvergiert für  $h \rightarrow 0$ . Die numerischen Lösungen sind vollständig von Oszillationen befreit und die Unstetigkeiten sind weitgehend erhalten. Die experimentelle Konvergenzordnung zeigt allerdings, daß das neue Verfahren nur von erster Ordnung genau ist.

## 4.3 Vollständig diskrete Filterroutinen

Eine weitere Möglichkeit, diskrete Filter für das numerische Lösen hyperbolischer Erhaltungsgleichungen zu entwickeln, ist es, den durch Differenzenverfahren höherer Ordnung verbrauchten Datensatz zu betrachten und durch rein diskrete Überlegungen und Operationen Algorithmen zu konstruieren, die Extremwerte vermindern.

Diese vollkommen diskrete Vorgehensweise haben auch Engquist, Lötstedt und Sjögreen zur Konstruktion ihrer, in Kapitel 2 vorgestellten, Algorithmen angewendet. Für einfache Anwendungen erzielen sie mit dieser diskreten Methode gute Filterverfahren für Differenzenschema höherer Ordnung.

Im folgenden Abschnitt konstruieren wir ebenfalls aus ausschließlich diskreten Betrachtungen eine Serie nichtlinearer Filterroutinen und zeigen ihre Anwendung auf klassische Problemstellungen.

### 4.3.1 Diskrete Filterentwicklung auf Grundlage der Algorithmen von Engquist, Lötstedt und Sjögreen

#### Herleitung der diskreten Filterroutine

Das diskrete Filtern eines durch Oszillationen verrauschten Datensatzes unterteilt sich in zwei Vorgänge, einerseits das Verringern der Extremwerte, das als Abschneiden einer bestimmten Fläche aufgefaßt werden kann, andererseits das neue Verteilen dieser entfernten Fläche auf den Datensatz, wodurch die Konservativität des resultierenden Filterverfahrens gewährleistet wird.

Den Reduktionsvorgang der Extremwerte übernehmen wir direkt von dem in Kapitel 2 vorgestellten Filteralgorithmus von Engquist, Lötstedt und Sjögreen. Sie verringern jeden Extremwert des Datensatzes um die Hälfte des kleinsten Abstandes zu seinen Nachbarwerten. Von lokalen Maxima wird die Hälfte der kleinsten Differenz zwischen Extremwert und Nachbarwert abgezogen, bei einem lokalen Minimum wird dieser Betrag addiert. Damit ist der erste Teilvorgang des zu konstruierenden Filters, das Abschneiden der Extremwerte, vollständig beschrieben.

In einem zweiten Schritt ist die von den Extremwerten abgetragene Fläche wieder auf den Datensatz zu verteilen. Bei dem Filter von Engquist und Mitarbeitern wird die am Extremwert an der Stelle  $i$  abgeschnittene Fläche auf einen der beiden direkten Nachbarn des Extremums, also auf den Wert an der Stelle  $i - 1$  oder  $i + 1$ , verteilt. Von diesem lokalen Flächenausgleich wollen wir bei der Konstruktion neuer diskreter Filterroutinen abweichen und eine globale Flächenverteilung wählen.

Zunächst addieren wir alle von den Extremwerten abgeschnittenen Flächen, wobei die Flächen, die von den Maxima abgetragen wurden, positives Vorzeichen und Flächen, die von den Minima abgetragen wurden, negatives Vorzeichen erhalten. Durch die Addition gleichen sich von oben und unten abgeschnittene Flächen bereits zu einem Großteil aus, so daß nur noch die resultierende Restfläche verteilt werden muß.

Zunächst beschreiben wir ein Verteilungsverfahren für die Restfläche, das in den späteren Unterkapiteln auf verschiedene Weisen modifiziert wird.

Eine erste Verteilungsmöglichkeit der Restfläche wurde in Zusammenarbeit mit Thomas Brox im Rahmen des DFG-Projekts SO 363/9-1 entwickelt.

Das Vorzeichen der Restfläche bestimmt die Menge der Punkte, auf die die Restfläche anteilig verteilt wird. Besitzt die nach der Addition übrige Restfläche negatives Vorzeichen, das ist der Fall, wenn die Minima stärker hochkorrigiert als die Maxima runterkorrigiert wurden, wird die Restfläche anteilig auf die Punkte verteilt, die bereits durch das Differenzenverfahren oder das Abschneiden der Extremwerte heruntergezogen wurden. Die Punkte, die also insgesamt mit negativer Diffusion aus dem Differenzenschritt und dem ersten Teil des Filtervorganges hervorgehen, werden weiter reduziert, bis die negative Restfläche vollständig verteilt ist und sich ein Gleichgewicht zwischen positiver und negativer Fläche einstellt, das die Konservativität des Filtervorganges garantiert. Die anteilige Verteilung der Restfläche wird dabei quantitativ wie folgt vorgenommen. Für alle Punkte  $i$ , für die die Gesamtdiffusion aus Differenzenverfahren und Abschneiden der Extrema negativ ist, berechnet sich die neue Diffusion aus dem Wert der alten Diffusion plus dem Wert, der sich aus der Multiplikation der

alten Diffusion mit dem Quotienten aus der negativen Restfläche und der gesamten negativen, also bei der Extremareduktion von oben abgeschnittenen, Fläche ergibt. Bezeichnet  $D(i)$  die Gesamtdiffusion aus dem Differenzenverfahren und dem Abschneiden der Extremwerte am Punkt  $i$ ,  $F_r$  die negative Restfläche und  $F_g$  die gesamte, von oben abgeschnittene (und daher positive) Fläche, um die die Extrema reduziert wurden, so berechnet sich die neue Diffusion  $D^{\text{neu}}$  an der Stelle  $i$  gemäß

$$D^{\text{neu}}(i) = D(i) + D(i) \cdot \frac{F_r}{F_g}$$

für alle  $i$  mit  $D(i) < 0$ .

Ist die Restfläche dagegen positiv, wird sie auf alle Punkte  $i$  mit  $D(i) > 0$  verteilt. Bezeichnet  $F_r$  die jetzt positive Restfläche und  $F_g$  die gesamte Fläche, um die die Minima abgeschnitten wurden, so läßt sich die neue Diffusion für alle Punkte  $i$  mit  $D(i) > 0$  ebenfalls durch

$$D^{\text{neu}}(i) = D(i) + D(i) \cdot \frac{F_r}{F_g}$$

berechnen. Damit ist der zweite Teil der Filterroutine, das Verteilen der Restfläche, vollständig beschrieben und ein erster, neuer diskreter Filter ist konstruiert.

Die entwickelte Filterroutine erfüllt allerdings noch nicht alle wünschenswerten Eigenschaften eines optimalen diskreten Filters. Werden beim Abschneiden der Extremwerte stets die aktualisierten Diffusionen der Nachbarpunkte betrachtet, ist der Filter beispielsweise nicht invariant zu seiner Durchlaufrichtung. Näherungsweise kann die Invarianz allerdings durch iteratives Abschneiden der Extremwerte erreicht werden. Verwenden wir beim Abschneidevorgang die ursprünglichen Diffusionen aus dem Differenzenverfahren, läßt sich auch exakte Invarianz gewährleisten.

Ein weitaus bedeutenderes Fehlverhalten des neuen Filters ist die Eigenschaft, daß bei der Verteilung der Restfläche neue Extremwerte entstehen können, was bei sinnvollen diskreten Filteralgorithmen natürlich nicht möglich sein darf.

Bevor wir auf Verbesserungsmöglichkeiten eingehen, zeigen wir zunächst numerische Ergebnisse, die sich mit der neu konstruierten Filterroutine erzielen lassen.

## Numerische Resultate

### Eindimensionale skalare Testfälle

Wir betrachten das eindimensionale Cauchy-Problem (1.49) für Burgers' Gleichung, das wir mit dem Lax-Wendroff-Verfahren als grundlegendes Differenzenverfahren lösen. Zusätzlich implementieren wir nach jedem Zeitschritt des Differenzenschemas zwei Iterationen mit dem konstruierten diskreten Filter, wobei wir die nicht invariante Abschneideroutine für die Extrema wählen. Bei einer CFL-Zahl von 0.5 ergibt sich auf einem äquidistanten Gitter aus 400 Punkten nach  $t = 2.5$  die in Abbildung 4.55 dargestellte Lösung. Der linke Teil der Abbildung zeigt die gefilterte Lax-Wendroff-Lösung, der rechte Teil stellt sie im Vergleich zum ungefilterten Lax-Wendroff-Verfahren dar.

Reduzieren wir die CFL-Zahl auf 0.1, vergrößern sich die Oszillationen des Lax-Wendroff-Verfahrens und die Leistungsfähigkeit des Filters wird noch deutlicher. Abbildung 4.56 zeigt

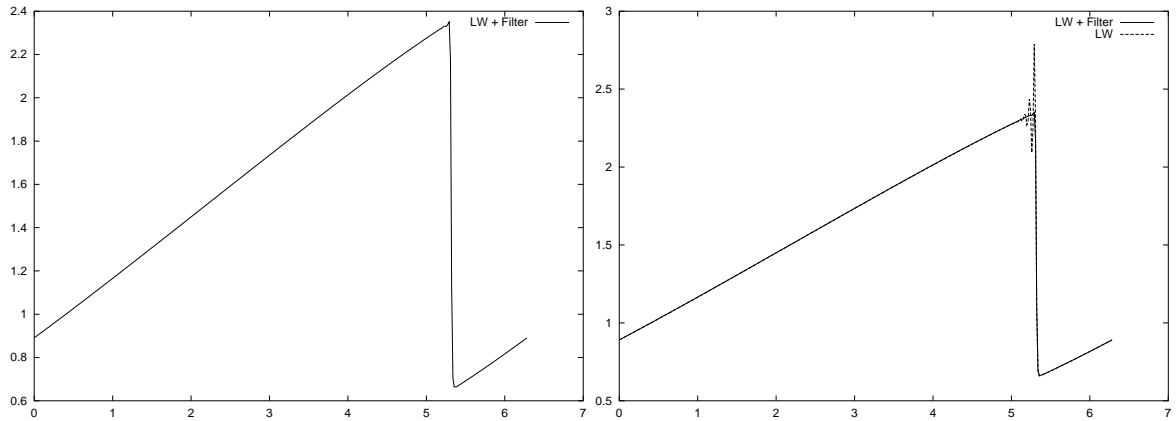


Abbildung 4.55: Gefilterte Lax-Wendroff-Lösung für das 1D-Cauchy-Problem (1.49) und  $CFL= 0.5$ , rechts im Vergleich zur reinen Lax-Wendroff-Lösung dargestellt

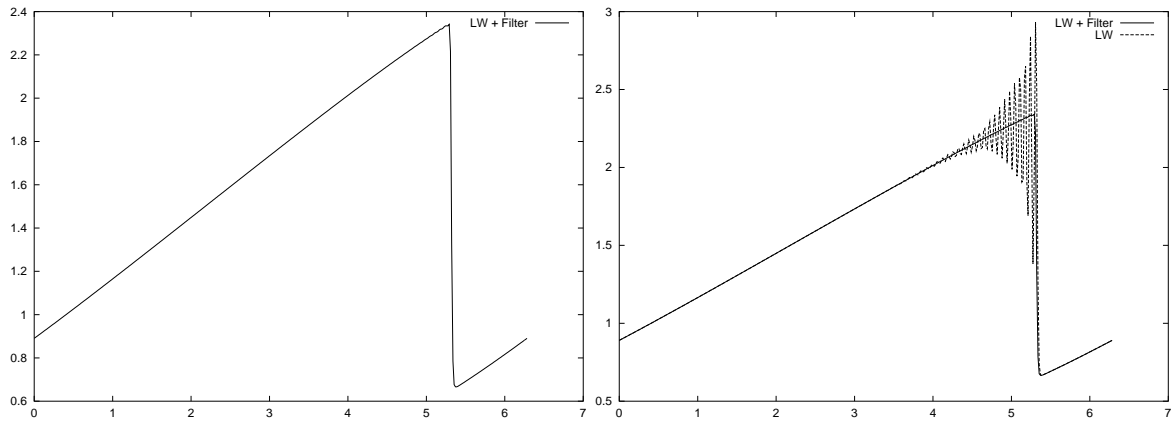


Abbildung 4.56: Gefilterte Lax-Wendroff-Lösung für das 1D-Cauchy-Problem (1.49) und  $CFL= 0.1$ , rechts im Vergleich zur reinen Lax-Wendroff-Lösung dargestellt

die mit nur einer Filteriteration pro Zeitschritt gefilterte Lax-Wendroff-Lösung für  $CFL= 0.1$  im Vergleich zur ungefilterten Lösung. Die Oszillationen des Differenzenverfahrens zweiter Ordnung sind geglättet, Lage und Höhe der Unstetigkeit sind erhalten, und das Verfahren konvergiert für alle  $CFL$ -Zahlen kleiner eins für  $h \rightarrow 0$ . Darüber hinaus ist das resultierende Verfahren sehr schnell, da eine iterative Entfernung der Extremwerte gar nicht nötig ist und schon ein einziger Filterdurchlauf pro Zeitschritt sehr gute Ergebnisse liefert. Probleme aufgrund der Invarianz oder des möglichen Entstehens neuer Extremwerte treten hier nicht auf. Das gleiche gilt, wenn wir das stabile Lax-Wendroff-Verfahren durch das instabile Zentrale-Differenzen-Schema zweiter Ordnung ersetzen und eine Iteration der nichtlinearen Filterroutine nach jedem Zeitschritt implementieren. Bei einer  $CFL$ -Zahl von 0.1 ergibt sich die im linken Teil von Abbildung 4.57 dargestellte Lösung. Auch hier sind die Oszillationen der zentralen Differenz geglättet und Lage und Höhe der Unstetigkeit sind erhalten. Allerdings sind, wie auch schon bei anderen, in das zentrale Verfahren implementierten Filtern, kleine Oszillationen niedriger Frequenz im linearen Anteil der Lösung zu erkennen, die auf die zentrale Differenz zurückzuführen sind. Für kleiner werdende Gitterweite  $h$  werden diese Störungen allerdings immer kleiner, wie das rechte Bild in Abbildung 4.57 für  $N = 1600$ , also viermal

so kleiner Raumschrittweite, zeigt, und das Verfahren konvergiert für  $h \rightarrow 0$ .

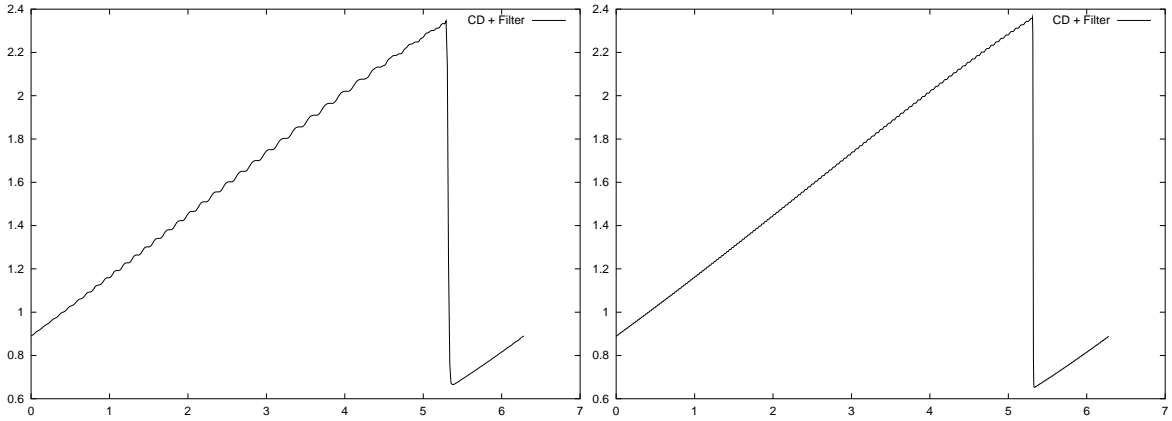


Abbildung 4.57: Gefilterte Zentrale-Differenzen-Lösung für das 1D-Cauchy-Problem (1.49) für  $N = 400$  (links) und  $N = 1600$  Gitterpunkte (rechts)

### Eindimensionale Systeme von Erhaltungsgleichungen

Für Anwendungen des diskreten Filters auf die eindimensionalen Eulergleichungen betrachten wir das einfache Riemann-Problem (1.58) ohne sonische Punkte mit einer Unstetigkeit an der Stelle  $x_0 = 0.3$ . Wir lösen das Anfangswertproblem mit dem Lax-Wendroff-Verfahren und implementieren nach jedem Zeitschritt eine Iteration des entwickelten Filters, indem wir jede Systemkomponente unabhängig voneinander filtern. Mit nur einer Filteriteration ergibt sich bei einer CFL-Zahl von 0.5 auf einem äquidistanten Gitter aus 400 Punkten nach  $t = 0.2$  die in Abbildung 4.58 links dargestellte gefilterte Lösung. Zum Vergleich ist die zugehörige ungefilterte Lax-Wendroff-Lösung dargestellt. Der rechte Teil der Abbildung zeigt das Resultat des gleichen numerischen Experiments, allerdings mit dem Zentrale-Differenzen-Verfahren zweiter Ordnung statt des Lax-Wendroff-Schemas.

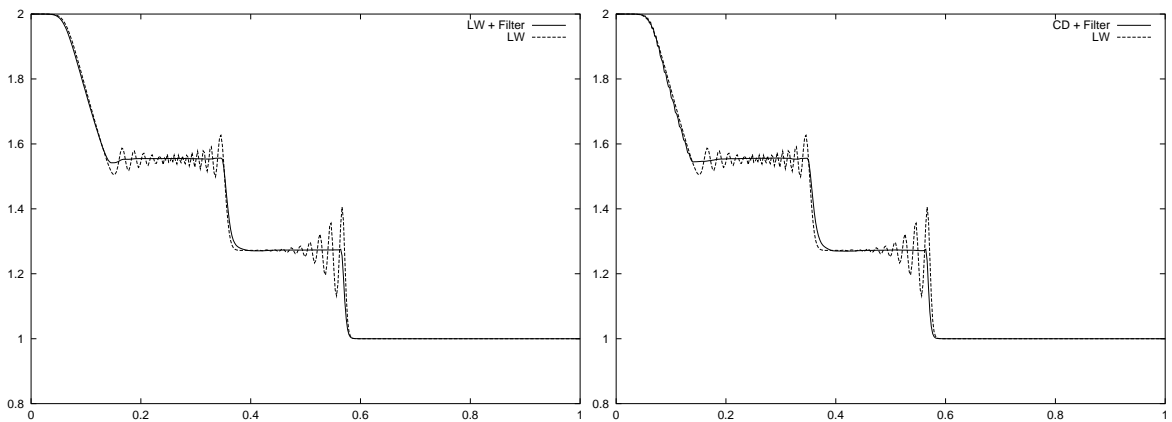


Abbildung 4.58: Gefilterte Lax-Wendroff- (links) und Zentrale-Differenzen-Lösung (rechts) für das Riemann-Problem (1.58)

In beiden gefilterten Lösungen sind die Oszillationen geglättet und Lage und Höhe der Unstetigkeiten sind erhalten. Die Filterroutine ist in der Lage, die instabile zentrale Differenz zu stabilisieren und führt selbst hier zu guten Ergebnissen. Beide Verfahren konvergieren für  $h \rightarrow 0$  und weisen keine Beeinträchtigung der Lösung aufgrund der Invarianz oder dem möglichen Entstehen neuer Extremwerte auf.

Diese Eigenschaften fallen erst bei der Filteranwendung auf Riemann-Probleme mit sonischen Punkten auf. Wir betrachten das Anfangswertproblem (1.60) mit einer Unstetigkeit bei  $x_0 = 0.5$ , das wir mit dem Lax-Wendroff-Verfahren, ergänzt durch eine Iteration des nichtlinearen Filters pro Zeitschritt, lösen. Unter den gleichen Bedingungen wie vorher ergibt sich nach  $t = 0.25$  die links in Abbildung 4.59 dargestellte gefilterte Lösung. Bis auf einen Einbruch der Daten etwa an der Stelle 0.5 sieht das Resultat, verglichen mit der ebenfalls dargestellten exakten Lösung, zufriedenstellend aus. Implementieren wir allerdings pro Zeitschritt zehn Filteriterationen des nichtlinearen Filters, ergibt sich das im rechten Teil der Abbildung dargestellte Resultat. Es ist zu erkennen, daß der entwickelte Filteralgorithmus

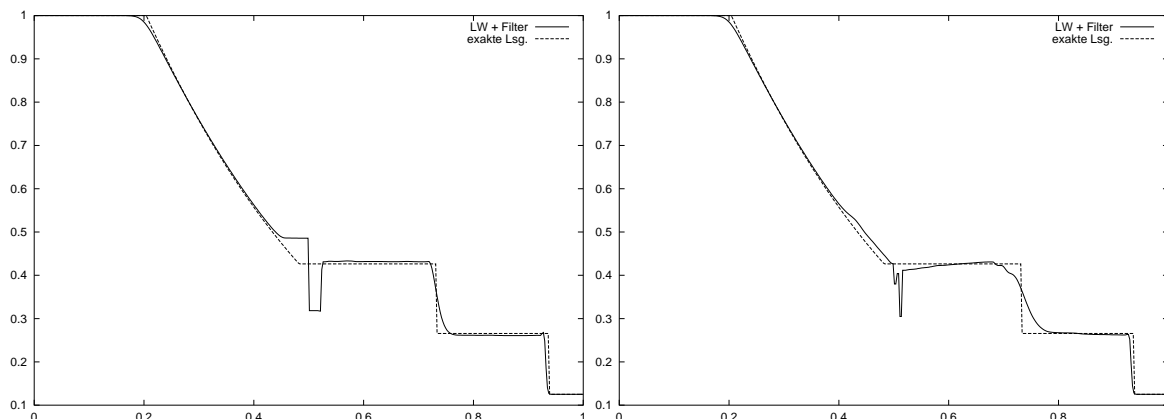


Abbildung 4.59: Gefilterte Lax-Wendroff-Lösung für (1.60) bei einer Filteriteration pro Zeitschritt (links) beziehungsweise zehn Iterationen (rechts)

bei steigender Iterationszahl versucht, das unphysikalische Minimum im Datensatz zu beseitigen. Zwar reduziert der Filter die Größe des Minimums tatsächlich, allerdings zerstört er beim Umverteilen der abgeschnittenen Fläche die Qualität der Lösung an anderen Stellen und neue Extremwerte, wenn auch erst kleine, entstehen. Sehr deutlich wird dieser Effekt, wenn wir die Simulation auf einem äquidistanten Gitter, bestehend aus doppelt so vielen Gitterpunkten, wiederholen. Das Lax-Wendroff-Verfahren mit zehn Filteriterationen nach jedem Zeitschritt wird in diesem Fall instabil und führt zu keiner Lösung. Die Eigenschaft des Filters, das Entstehen neuer Extremwerte nicht zu unterbinden, zerstört die Lösung in diesem Fall vollständig. Implementieren wir dagegen nur eine Filteriteration nach jedem Zeitschritt des Lax-Wendroff-Verfahrens, bleibt das resultierende Lösungsverfahren auch für kleiner werdende Gitterweite stabil. Für das Riemann-Problem (1.58) ohne sonische Punkte bleibt die Lösung auch bei zehn Filteriterationen pro Zeitschritt stabil und konvergiert für  $h \rightarrow 0$ . Dennoch zeigt die Filteranwendung auf das Sod-Problem (1.60), speziell bei der Erhöhung der Iterationszahl, daß eine weitere Verbesserung des Filters notwendig ist. Mögliche Variationen werden im Folgenden vorgestellt und auf ihre numerische Anwendung hin überprüft.



### 4.3.2 Erste Filtervariante

Wie das letzte Anwendungsbeispiel zeigt, stellt die Verteilung der Restfläche bei der entwickelten Filterroutine ein Problem dar, so daß wir an dieser Stelle Korrekturen vornehmen.

#### Alternative Restflächenverteilung

Da das zugrundeliegende Differenzenverfahren den Datensatz, abgesehen von den Oszillationen, die entstehen, physikalisch sinnvoll ändert, ist es naheliegend, die nach Ausgleich übrige Restfläche nicht auf alle Punkte  $i$  mit  $D(i) < 0$  beziehungsweise  $D(i) > 0$  zu verteilen, sondern lediglich auf die entstanden Extremwerte, also auf die Punkte, die bereits beim Abschneidevorgang berücksichtigt wurden. Die anteilige Verteilung der Restfläche wird dabei genauso wie vorher vorgenommen. Ist die Restfläche negativ und bezeichnet  $D(i)$  jetzt die Diffusion am Punkt  $i$ , die lediglich aus dem Abtragen der Extremwerte hervorgeht, so ist der Verteilungsalgorithmus wieder durch

$$D^{\text{neu}}(i) = D(i) + D(i) \cdot \frac{F_r}{F_g}$$

für alle  $i$  mit  $D(i) < 0$  gegeben, wobei  $F_r$  die negative Restfläche und  $F_g$  die gesamte, von oben abgeschnittene Fläche bezeichnet. Entsprechendes gilt, wenn die Restfläche positiv ist.

#### Numerische Anwendung

Implementieren wir diese Flächenverteilung in die Filterroutine, ohne weitere Änderungen vorzunehmen, müssen wir feststellen, daß schon bei der Anwendung auf das einfache Cauchy-Problem für Burgers' Gleichung, das mit dem Lax-Wendroff-Verfahren und einer Iteration der neuen Filtervariante nach jedem Zeitschritt gelöst wird, ein Verschiebungsfehler in der Lösung auftritt. Abbildung 4.60 zeigt die gefilterte Lösung im Vergleich zu dem entsprechenden Resultat mit der ursprünglichen Filterroutine.

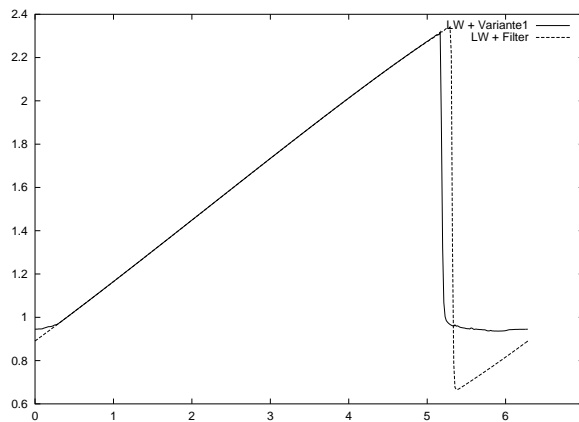


Abbildung 4.60: Vergleich der Filterwirkung vom ursprünglichen und abgewandelten Filteralgorithmus

Die Verschiebung der Unstetigkeit bei Anwendung der abgeänderten Filterroutine ist dadurch zu erklären, daß die Restfläche nicht mehr gleichmäßig auf die Daten verteilt wird, sondern hauptsächlich auf die Punkte, die hinter der Unstetigkeit liegen. Dadurch wird das Wertenniveau dort künstlich angehoben, wie in Abbildung 4.60 zu erkennen ist, und die Stoßausbreitungsgeschwindigkeit, die abhängig von der Stoßhöhe ist, verändert sich und verursacht eine Verschiebung der Stoßlage.

Damit ist gezeigt, daß diese Art der Flächenverteilung zu keinen sinnvollen Ergebnissen führt.

Zwar kann die Restflächenverteilung weiterhin so variiert werden, daß die Restfläche nicht nur auf die reduzierten Extrema, sondern auch auf die beiden jeweils direkten Nachbarn verteilt wird, so daß kein Verschiebungsfehler mehr auftritt, aber da sich die Ergebnisse nicht sehr von denen des ersten Filters unterscheiden, verzichten wir auf nähere Betrachtungen.

### 4.3.3 Zweite Variante: Behebung der Filtermängel

#### Entwicklung der neuen Filterroutine

In einem zweiten Versuch, die konstruierte Filterroutine sinnvoll zu korrigieren, ändern wir zunächst die Invarianzeigenschaft des Filters. Die bisher fehlende Invarianz zur Durchlaufrichtung gewährleisten wir dadurch, daß beim Abschneiden der Extremwerte stets der jeweilige Ausgangssatzenz, der sich nach dem Zeitschritt des Differenzschemas ergibt, verwendet wird.

Um das Auftreten neuer Extremwerte bei der Verteilung der nach Addition übrigen Restfläche zu verhindern, verteilen wir eine positive Restfläche auf die vorher erkannten und reduzierten Maxima beziehungsweise eine negative Restfläche auf die Minima des Datensatzes. Das heißt, entweder werden die Maxima, die reduziert wurden, wieder etwas erhöht, oder die Minima, die erhöht wurden, werden wieder leicht verringert. Dadurch wird der Filtereffekt zwar etwas verringert, aber dafür entstehen keine neuen Extremwerte. Probleme treten bei dieser Verteilung allerdings auf, wenn in dem Datensatz nur Minima oder nur Maxima existieren, da dann kein Flächenausgleich stattfinden kann und der Filter nichts bewirkt. Da die Oszillationen, die durch ein Differenzenverfahren höherer Ordnung hervorgerufen werden, in der Regel stets gleichmäßig als Minima und Maxima verteilt auftreten, stellt diese Eigenschaft in der Praxis kein Problem dar.

Die Flächenverteilung auf die Minima oder Maxima findet wieder, wie im ursprünglichen Algorithmus, anteilig statt. Ist die Restfläche negativ und bezeichnet  $D(i)$  die Diffusion am Punkt  $i$ , die lediglich aus dem Filterschritt, also dem Abtragen der Extremwerte hervorgeht, so berechnet sich die neue Diffusion  $D^{\text{neu}}(i)$ , die durch die Verteilung der Restfläche entsteht, an diesem Punkt gemäß

$$D^{\text{neu}}(i) = D(i) + D(i) \cdot \frac{F_r}{F_g},$$

und dies gilt für alle  $i$  mit  $D(i) < 0$ .  $F_r$  bezeichnet dabei die Restfläche und  $F_g$  stellt nun die gesamte Fläche, die von unten, also von den Minima, abgeschnitten wurde, dar.

Ist die Restfläche positiv, berechnet sich die neue Diffusion für alle Punkte  $i$  mit  $D(i) > 0$  aus der gleichen Formel, wobei  $F_g$  dann die gesamte Fläche, die von oben, also von den Maxima, abgetragen wurde, bezeichnet.

Damit ist eine neue Filterroutine, die invariant zur Durchlaufrichtung der Daten ist und keine neuen Extremwerte erzeugt, konstruiert und kann auf die verschiedenen Testfälle angewendet werden.

## Numerische Resultate

Für eine gute Vergleichbarkeit der numerischen Resultate wiederholen wir die bereits für den ursprünglichen Filter durchgeführten numerischen Simulationen unter gleichen Bedingungen.

### Eindimensionale skalare Testfälle

Das eindimensionale Cauchy-Problem (1.49) für Burgers' Gleichung, gelöst mit dem Lax-Wendroff-Verfahren und einer Filteriteration nach jedem Zeitschritt, ergibt auf dem äquidistanten Gitter aus 400 Punkten bei einer CFL-Zahl von 0.1 nach  $t = 2.5$  die im linken Teil von Abbildung 4.61 dargestellte Lösung. Direkt vor der Unstetigkeit ist ein großer Peak zu erkennen, den der Filter aufgrund seiner Restflächenverteilung nicht glätten kann. Außerdem sind im linearen Teil der Lösung sehr kleine Oszillationen zu erkennen, die ebenfalls dadurch entstehen, daß erkannte Maxima bei der Flächenumverteilung wieder leicht erhöht oder Minima wieder leicht reduziert werden. Wird die Gitterweite halbiert, ist dieser Effekt bereits nicht mehr sichtbar und auch der Peak vor der Unstetigkeit wird schmaler, wie der rechte Teil der Abbildung 4.61 für  $N = 800$  zeigt. Für  $h \rightarrow 0$  konvergiert das Verfahren. Gleiches

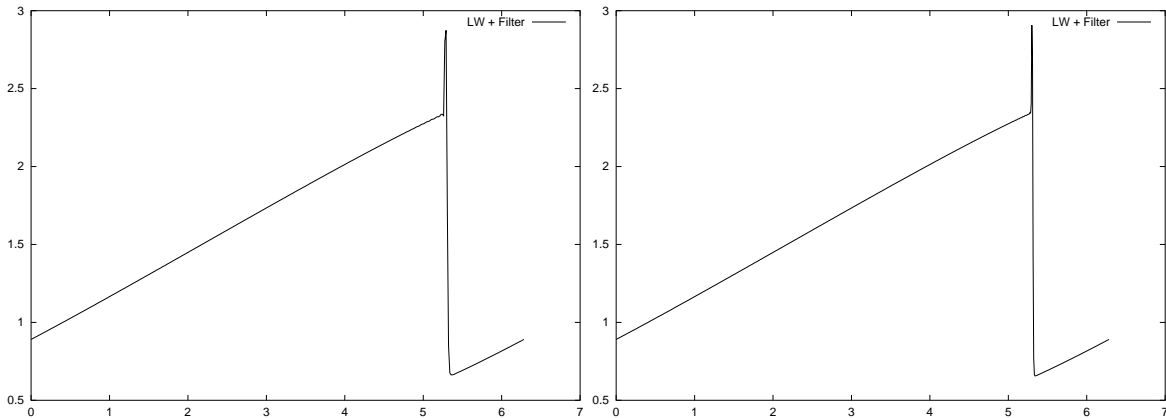


Abbildung 4.61: Gefilterte Lax-Wendroff-Lösung auf einem Gitter aus  $N = 400$  (links) und  $N = 800$  (rechts) äquidistanten Gitterpunkten

gilt für alle CFL-Zahlen kleiner eins. Für die dargestellte Simulation wurde sie lediglich auf 0.1 gesetzt, da das Lax-Wendroff-Verfahren dann größere Oszillationen erzeugt und die Leistungsfähigkeit des Filters deutlicher zu erkennen ist. Ersetzen wir das Lax-Wendroff-Verfahren durch das instabile Zentrale-Differenzen-Verfahren zweiter Ordnung, ergeben sich unter gleichen Bedingungen auf einem Gitter aus 400 beziehungsweise 1600 äquidistanten Punkten die in Abbildung 4.62 dargestellten Ergebnisse. Die Lösung auf 400 Gitterpunkten weist im linearen Teil noch kleine Oszillationen niedriger Frequenz auf, die aus dem Zentrale-Differenzen-Verfahren resultieren. Für kleiner werdende Gitterweite treten sie nicht mehr auf, wie der rechte Teil von Abbildung 4.62 für  $N = 1600$  andeutet.

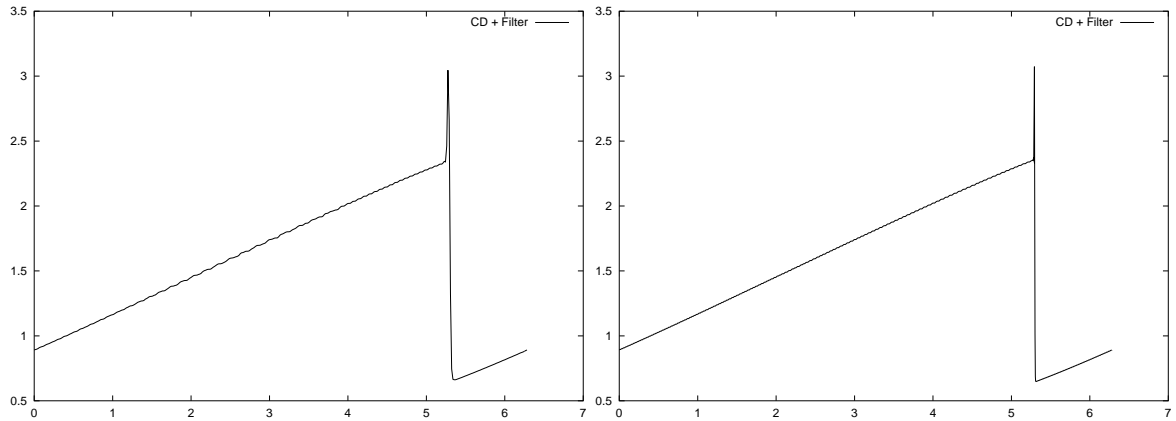


Abbildung 4.62: Gefilterte Zentrale Differenzen-Lösung auf einem Gitter aus  $N = 400$  (links) und  $N = 1600$  (rechts) äquidistanten Gitterpunkten

### Eindimensionale Systeme von Erhaltungsgleichungen

Für das einfache Riemann-Problem (1.58) ohne sonische Punkte, das mit dem Lax-Wendroff- beziehungsweise dem Zentrale-Differenzen-Verfahren und einem Filterschritt nach jedem Zeitschritt gelöst wird, ergeben sich die in Abbildung 4.63 gezeigten Ergebnisse. Abgesehen

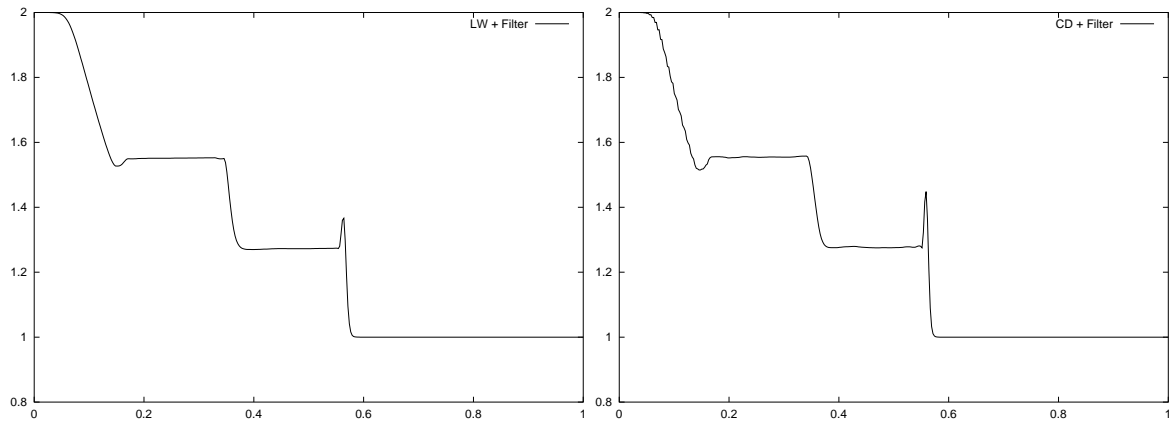


Abbildung 4.63: Gefilterte Lax-Wendroff- (links) und Zentrale-Differenzen-Lösung (rechts) für das Riemann-Problem (1.58)

von dem Peak vor der letzten Unstetigkeit, der sowohl beim gefilterten Lax-Wendroff- als auch beim gefilterten Zentrale-Differenzen-Verfahren auftritt, ist das Ergebnis besonders im Falle der zentralen Differenz beeindruckend, zumal nur eine Filteriteration pro Zeitschritt verwendet wird und die CFL-Zahl 0.5 beträgt. Ohne Filterroutine wird das Zentrale-Differenzen-Schema unter diesen Bedingungen fast sofort instabil. Sowohl das gefilterte Lax-Wendroff- als auch das gefilterte Zentrale-Differenzen-Verfahren konvergieren für  $h \rightarrow 0$ . Eine künstliche Verkleinerung der hyperbolischen Zeitschrittweite ist nicht notwendig, so daß die resultierenden Verfahren nur wenig Rechenzeit benötigen.

Tabelle 4.4 zeigt die Fehlernorm und die experimentelle Konvergenzordnung des mit einer

Filteriteration pro Zeitschritt gefilterten Lax–Wendroff–Verfahrens bei Anwendung auf das Riemann–Problem (1.58) in Abhängigkeit der Gitterweite  $h$ . Bleibt an den Unstetigkeitsstellen jeweils ein Intervall der Länge  $\Delta I = 0.023$  bei der Berechnung unberücksichtigt, hat das resultierende Verfahren, außer für  $h = 0.01$ , eine experimentelle Konvergenzordnung größer eins und ist damit ein Verfahren zweiter Ordnung. Die entsprechenden Werte für das mit einer

h	LW + 1 Filterit. $\ u - u_h\ _{L_1}$ EOC	LW + 1 Filterit. ohne Unstetigk. $\ u - u_h\ _{L_1}$ EOC
0.02	0.032017	0.023337
0.01	0.017995 0.83	0.013504 0.79
0.005	0.010748 0.74	0.006609 1.04
0.0025	0.005831 0.88	0.002916 1.19
0.00125	0.002955 0.98	0.001219 1.28

Tabelle 4.4: Experimentelle Konvergenzordnung des mit einer Filteriteration pro Zeitschritt gefilterten Lax–Wendroff–Verfahrens

h	CD + 1 Filterit. $\ u - u_h\ _{L_1}$ EOC	CD + 1 Filterit. ohne Unstetigk. $\ u - u_h\ _{L_1}$ EOC
0.02	0.036872	0.028724
0.01	0.020085 0.88	0.015507 0.92
0.005	0.013223 0.60	0.008739 0.86
0.0025	0.008440 0.65	0.004556 0.99
0.00125	0.005616 0.59	0.002773 0.72

Tabelle 4.5: Experimentelle Konvergenzordnung des mit einer Filteriteration pro Zeitschritt gefilterten Zentrale–Differenzen–Verfahrens

Filteriteration pro Zeitschritt gefilterte Zentrale–Differenzen–Schema sind weitaus schlechter, und selbst beim Aussparen der Unstetigkeiten ergibt sich nur eine Konvergenzordnung kleiner eins, wie Tabelle 4.5 zeigt.

Für das Riemann–Problem (1.60) von Sod, bei dem ein sonischer Punkt entsteht, ergibt das Lax–Wendroff–Verfahren mit einer Filteriteration nach jedem Zeitschritt für CFL= 0.5 und 400 Gitterpunkte das in Abbildung 4.64 links dargestellte Resultat. Wie bei der Verwendung des ursprünglichen Filters ist das in diesem Fall instabile Differenzenverfahren stabilisiert, und es ergibt sich eine stationäre Lösung. Dennoch ist ein Dateneinbruch in der Nähe des sonischen Punktes zu erkennen und die horizontalen Werteniveaus werden nicht exakt nachgebildet, wie ein Vergleich mit der exakten Lösung zeigt. Implementieren wir zehn Filteriterationen nach jedem Zeitschritt des Lax–Wendroff–Verfahrens, bleibt der Dateneinbruch, im Gegensatz zur ursprünglichen Filterroutine, konstant erhalten, aber trotzdem wird die Lösung an anderen Stellen schlechter, wie der rechte Teil der Abbildung verdeutlicht. Für kleiner werdende Gitterweite wird dieser Effekt allerdings wieder kleiner, wie Abbildung 4.65 für  $N = 1600$  Gitterpunkte zeigt.

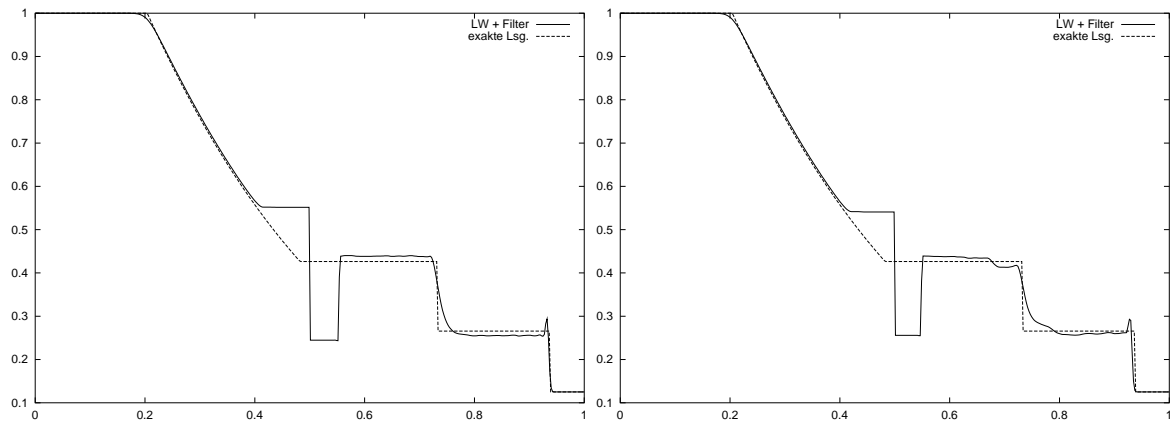


Abbildung 4.64: Gefilterte Lax–Wendroff–Lösung für das Sod–Problem (1.60) mit einer (links) und zehn (rechts) Filteriterationen pro Zeitschritt

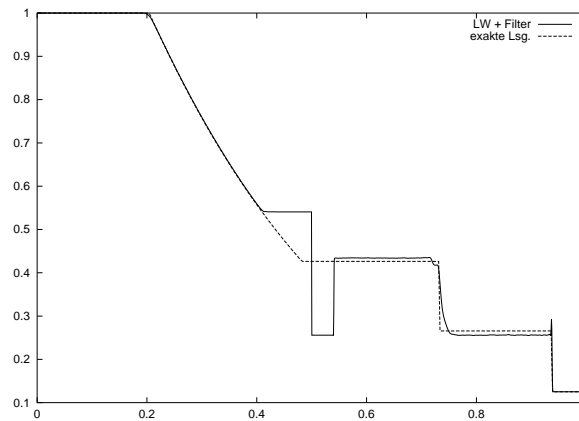


Abbildung 4.65: Gefilterte Lax–Wendroff–Lösung mit zehn Filteriterationen pro Zeitschritt und  $N = 1600$

Bei dem ursprünglichen Filter kann diese Verbesserung bei kleinerer Gitterweite nicht beobachtet werden, im Gegenteil, schon für  $N = 800$  wird das Verfahren instabil.

Dennoch wird deutlich, daß auch diese Filtervariante nicht stark genug ist, um unphysikalische Störungen an sonischen Punkten vollständig zu vermeiden.

### Zusätzliche Hintergrunddiffusion

Die durchgeführten Experimente zeigen, daß Riemann–Probleme mit sonischen Punkten ein besonderes Problem für die Differenzenverfahren mit integriertem Filter darstellen, da die an sonischen Punkten auftretenden Oszillationen weitaus größer sind als die typischen Oszillationen von Differenzenverfahren höherer Ordnung. Daher ist für diese Anwendungsprobleme eine stärkere Glättung als die der konstruierten Filterroutine notwendig. Ein naheliegendes Verfahren, das Abhilfe schafft, ist die zusätzliche Addition einer Hintergrunddiffusion in das Differenzenschema. Um eine geeignete Diffusion zu finden, orientieren wir uns an Kroll und Jain [31], die eine Hintergrunddiffusion, die sich aus zweiten und vierten Differenzen zusammensetzt, konstruieren. Wir übernehmen aus dieser Diffusion lediglich letzteren Term,

einen Ausdruck der Form  $\varepsilon^{(4)} u_{xxxx}$ , da die zweite Differenz für unsere Anwendungen zu stark glättend wirkt.

Bezeichnet  $D_i$  die Hintergrunddiffusion, so ergibt ihre Addition zum Differenzenverfahren das Differenzenschema

$$U_i^{n+1} = U_i^n - \lambda(F_{i+1/2}^n - F_{i-1/2}^n) + D_i.$$

Für die Diffusion von Kroll und Jain gilt

$$D_i = d_{i+1/2} - d_{i-1/2}$$

mit

$$d_{i+1/2} := -\alpha_{i+1/2} \varepsilon_{i+1/2}^{(4)} (U_{i+2} - 3U_{i+1} + 3U_i - U_{i-1}),$$

wobei die zweite Differenz an dieser Stelle bereits weggelassen ist. Der Koeffizient  $\alpha_{i+1/2}$  dient der richtigen Gewichtung des Diffusionsterms und kann in unserem Fall gemäß

$$\alpha_{i+1/2} = \frac{h}{k}$$

berechnet werden und ist damit unabhängig vom Gitterpunkt  $i$ . Auch  $\varepsilon_{i+1/2}^{(4)}$  ist bei Kroll und Jain ein adaptiver Koeffizient, der dazu dient, die Größe der Diffusion zu steuern. Für unsere Anwendung reduziert er sich auf eine Konstante  $K^{(4)}$ , für die nach umfassenden numerischen Experimenten von Schmidt und Jameson [57] beziehungsweise Jameson und Baker [28] die Größenordnung

$$\frac{1}{128} \leq K^{(4)} \leq \frac{1}{64}$$

festgelegt wurde. Wir wählen für unsere Anwendungen den größten Wert für  $K^{(4)}$ , also  $K^{(4)} = \frac{1}{64}$ . Die Hintergrunddiffusion ist damit vollständig beschrieben und kann in das Differenzenschema, in unserem Fall das Lax–Wendroff–Verfahren, implementiert werden.

Zunächst versuchen wir, das Sod–Problem (1.60) lediglich mit der Hintergrunddiffusion im Lax–Wendroff–Schema bei einer CFL–Zahl von 0.5 auf einem Gitter aus 400 äquidistanten Punkten zu lösen, um festzustellen, welche Auswirkung allein die zusätzliche Diffusion auf das Differenzenverfahren hat. Dabei wird deutlich, daß die zusätzliche Diffusion allein nicht ausreicht, um das Lax–Wendroff–Verfahren am sonischen Punkt zu stabilisieren. Abbildung 4.66 links zeigt die resultierende Lösung nach der Zeit  $t = 0.01$ , die starke Oszillationen aufweist, und wenig später bricht das Verfahren vollständig zusammen. Implementieren wir dagegen nach jedem Zeitschritt des Lax–Wendroff–Verfahrens mit integrierter Diffusion eine Iteration der zuletzt hergeleiteten Filterroutine, ergibt sich unter gleichen Bedingungen die im rechten Teil von Abbildung 4.66 dargestellte Lösung. Die bisherigen Probleme in der Umgebung des sonischen Punktes sind bis auf einen kleinen Unterschwinger behoben, und auch die Qualität der gesamten Lösung ist im Vergleich zu vorherigen Lösungen des Sod–Problems (1.60) gestiegen. Das Zusammenspiel aus Hintergrunddiffusion und Filterroutine liefert ein stabiles Lösungsverfahren, das für  $h \rightarrow 0$  konvergiert.

Erhöhen wir die Anzahl der Filteriterationen pro Zeitschritt beispielsweise auf 10, wird das Ergebnis zwar zunächst etwas schlechter, wie im linken Teil der Abbildung 4.67 zu erkennen ist, aber eine Verkleinerung der Gitterweite hebt diesen Effekt wieder auf, wie die Lösung, berechnet auf einem Gitter aus 1600 statt 400 Punkten, im rechten Teil der Abbildung zeigt. Das konstruierte Verfahren konvergiert für  $h \rightarrow 0$  auch bei größerer Anzahl von Filteritera-

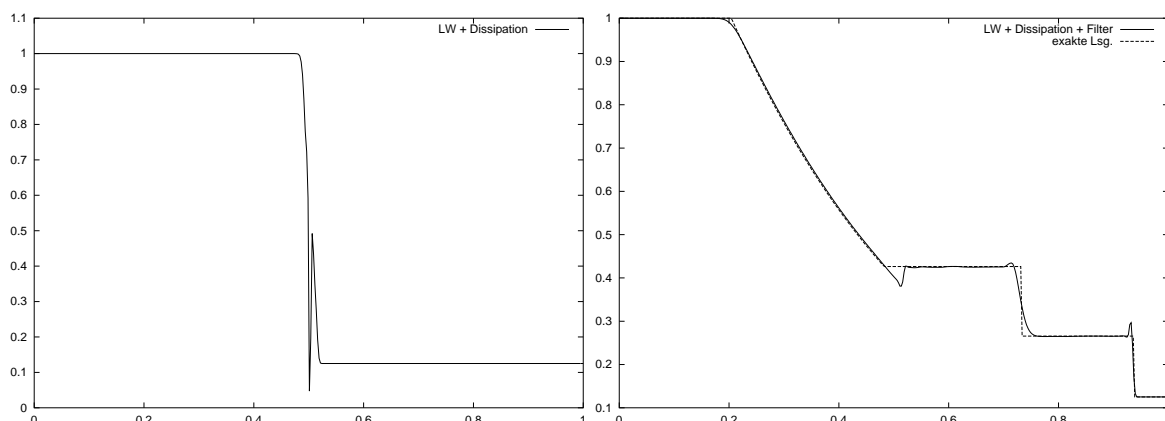


Abbildung 4.66: Lax–Wendroff–Lösung mit integrierter Hintergrunddiffusion ohne Filter zur Zeit  $t = 0.01$  (links) beziehungsweise mit einer Filteriteration zur Zeit  $t = 0.25$  (rechts)

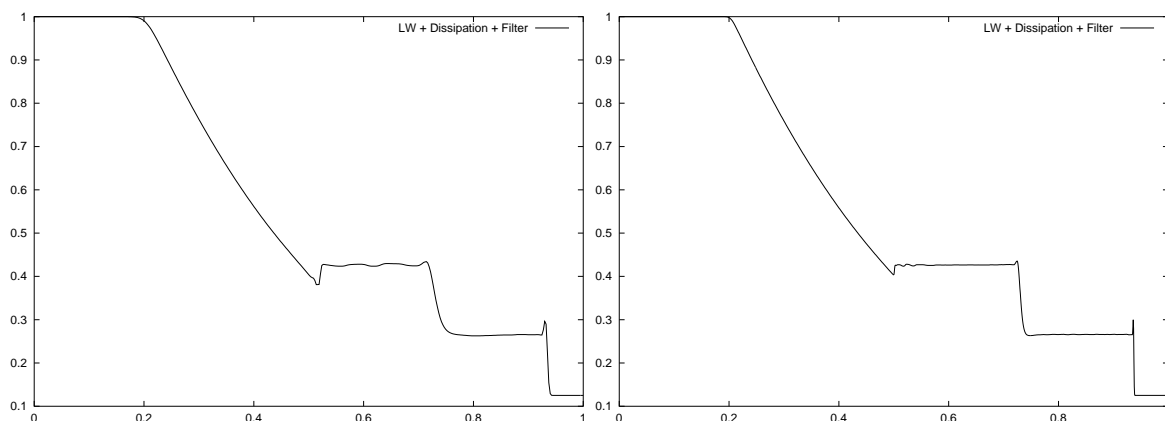


Abbildung 4.67: Mit 10 Iterationen gefilterte Lax–Wendroff–Lösung mit integrierter Hintergrunddiffusion für  $N = 400$  (links) und  $N = 1600$  (rechts) Gitterpunkte

tionen pro Zeitschritt.

Da die Hintergrunddiffusion die Größenordnung der Rechengeschwindigkeit nicht beeinflusst, ist das resultierende Verfahren bei einer Filteriteration pro Zeitschritt von geringem rechen-technischen Aufwand.

Ersetzen wir das Lax–Wendroff–Verfahren durch das Zentrale–Differenzen–Verfahren zweiter Ordnung, ergibt sich für das Sod–Problem (1.60) bei zusätzlicher Implementierung der Hintergrunddiffusion ab zwei Filteriteration pro Zeitschritt eine Lösung, die allerdings für kleiner werdende Gitterweite instabil wird.

#### 4.3.4 Zusammenfassung der Ergebnisse

Wir haben allein aus rein diskreten Überlegungen eine Reihe nichtlinearer, diskreter Filter entwickelt, die für das numerische Lösen eindimensionaler skalarer und vektorieller Anfangswertprobleme einsetzbar sind und teilweise zu guten numerischen Ergebnissen führen.

Schon der erste Filteralgorithmus, der noch nicht invariant zur Durchlaufrichtung ist und das



Entstehen neuer Extremwerte erlaubt, liefert für skalare Anwendungen sowohl für das Lax–Wendroff– als auch für das Zentrale–Differenzen–Verfahren sehr glatte Lösungen, in denen Höhe und Lage der Unstetigkeiten erhalten ist. Gleiches gilt für die Riemann–Probleme ohne sonische Punkte. Erst bei der Filteranwendung auf Riemann–Probleme mit sonischen Punkten zeigt sich, daß das resultierende Verfahren bei steigender Anzahl von Filteriterationen und kleiner werdender Gitterweite instabil wird.

Abhilfe schafft die Entwicklung der zweiten Filtervariante, die das Entstehen neuer Extremwerte verhindert und invariant zur Durchlaufzeit ist. Bei der Anwendung auf skalare Testfälle bleibt in der numerischen Lösung stets eine Oszillationsspitze direkt vor der Unstetigkeit erhalten, die durch die Filterstruktur zu erklären ist. Davon abgesehen sind die Lösungen sehr glatt und erhalten die Unstetigkeiten in Höhe und Lage. Die Berechnung der experimentellen Konvergenzordnung für die Lösung des Riemann–Problems (1.58) zeigt, daß die Integration einer Filteriteration nach jedem Zeitschritt des Lax–Wendroff–Schemas ein Verfahren zweiter Ordnung liefert. Bei der Implementierung in das Zentrale–Differenzen–Verfahren ergibt sich dagegen nur ein Verfahren erster Ordnung.

Für Riemann–Probleme mit sonischen Punkten liefert die Filterroutine, integriert nach jedem Zeitschritt des Lax–Wendroff–Verfahrens, auch für steigende Anzahl von Filteriterationen und  $h \rightarrow 0$ , stabile Lösungen, die allerdings einen Dateneinbruch am sonischen Punkt aufweisen. Doch mit Hilfe einer zusätzlichen Hintergrunddiffusion, die allein nicht leistungsfähig genug ist, um das Lax–Wendroff–Verfahren für Riemann–Probleme mit sonischen Punkten zu stabilisieren, ergibt die Filteranwendung selbst für diese Testprobleme gute numerische Resultate. Darüber hinaus liefern alle diskret entwickelten Filter sehr schnelle Lösungsverfahren, und Probleme, die sich bei der Diskretisierung kontinuierlicher Filtergleichungen ergeben, treten bei dieser vollständig diskreten Filterkonstruktion gar nicht erst auf.

Auch eine Verschiebung der Stoßlage, die wir bei der Implementierung einiger Diskretisierungen kontinuierlicher Filtergleichungen beobachtet haben, ist bei der diskreten Filterentwicklung nicht festzustellen, wenn die Restfläche gleichmäßig auf die Daten verteilt wird.

Damit bietet die vollständig diskrete Filterkonstruktion eine gute Alternative zur Herleitung diskreter Filter auf der Grundlage kontinuierlicher Gleichungen.



## 5 Zusammenfassung und Ausblick

In der vorliegenden Arbeit beschäftigen wir uns mit der Entwicklung nichtlinearer, diskreter Filteralgorithmen für das numerische Lösen hyperbolischer Erhaltungsgleichungen. Unser Ziel ist es, einfache Differenzenverfahren höherer Ordnung, die in ihren numerischen Lösungen Oszillationen in der Umgebung von Unstetigkeiten erzeugen, durch diskrete Filterroutinen so zu ergänzen, daß die Oszillationen bei gleichzeitiger Erhaltung der Unstetigkeiten geglättet werden.

Dem neuen Ansatz von Grahs [21] und Wei [74] folgend, basiert die Konstruktion unserer Filter einerseits auf Filteralgorithmen, die in der Bildverarbeitung entwickelt wurden und zum Wiederherstellen verrauschter Bilder eingesetzt werden. Andererseits beschäftigen wir uns mit vollständig diskreten Filterkonstruktionen, wie sie erstmals von Engquist, Lötstedt und Sjögreen für numerische Lösungen von Erhaltungsgleichungen durchgeführt wurden. Mit der Entwicklung einer kontinuierlichen Gleichung für einen ihrer diskreten Filter können wir einen Zusammenhang zwischen der vollständig diskreten Filterkonstruktion und den klassischen Filterherleitungen der Bildverarbeitung aufzeigen.

Für zwei bekannte TV-Filter der Bildverarbeitung zeigen wir, daß sie auf den Kontext numerischer Approximationen von Erhaltungsgleichungen übertragen und als nichtlineare Dissipation in einfache Differenzenverfahren höherer Ordnung integriert werden können.

Für den TV-Filter von Chan, Osher und Shen demonstrieren zweidimensionale skalare und vektorielle Testfälle, daß er bei Anwendung auf stationäre Lösungen Oszillationen in der Umgebung von Unstetigkeiten reduziert und Höhe und Lage der Unstetigkeiten erhält. Die zweidimensionalen Testfälle diskretisieren wir dabei auf unstrukturierten Gittern. Eindimensionale skalare Testfälle zeigen, daß der Filter, implementiert nach jedem Zeitschritt eines Differenzenverfahrens, selbst das instabile Zentrale-Differenzen-Verfahren zweiter Ordnung stabilisiert und eine von Oszillationen befreite Lösung erzeugt. Allerdings zeigt sich auch, daß die Lage der Unstetigkeit in der numerischen Lösung gegenüber der tatsächlichen Lage, abhängig von der Wahl der Filterparameter, verschoben dargestellt wird. Nähere Betrachtungen zeigen, daß eine Ursache für diese Verschiebung der Stoßlage die Tatsache ist, daß der diskrete Filter nur für wenige Parameterkombinationen konservativ ist und sich damit für die Filterimplementierung nach jedem Zeitschritt eines Differenzenverfahrens nicht eignet.

Die numerischen Resultate für den TV-Filter von Marquina und Osher, der auf dem gleichen kontinuierlichen Modell wie der Filter von Chan, Osher und Shen beruht, zeigen, daß seine Implementierung nach jedem Zeitschritt eines Differenzenverfahrens höherer Ordnung im Wesentlichen keine Verschiebung der Stoßlage bewirkt. Die Oszillationen der zugrunde-

liegenden Differenzenverfahren werden durch die Filteranwendung geglättet, und selbst das Zentrale-Differenzen-Schema zweiter Ordnung wird durch die Filterroutine stabilisiert und ergibt numerische Lösungen ohne Oszillationen.

Das unterschiedliche numerische Verhalten der diskreten Filter zeigt, daß Eigenschaften der kontinuierlichen Filtergleichung bei der Diskretisierung verloren gehen können und die Wirkungsweise diskreter Filter stark von der gewählten Diskretisierung abhängt.

Um dies zu belegen, entwickeln wir für die kontinuierliche Filtergleichung eine Reihe neuer Diskretisierungen und zeigen die Leistungsfähigkeit der resultierenden Filteralgorithmen für eindimensionale skalare Testfälle. Wir erhalten unter anderem einen parameterfreien Filter, der als einfache nichtlineare Dissipation das Zentrale-Differenzen-Verfahren stabilisiert, die Oszillationen glättet und Lage und Höhe der Unstetigkeit erhält. Eine Untersuchung der Gleichung  $u_t = (\operatorname{sgn} u_x)_x$ , die einerseits den neuen Filterroutinen zugrundeliegt, andererseits im Kontext der Bildverarbeitung als *TV flow* bekannt ist und als TV-verringern gilt, zeigt, daß die verschiedenen Diskretisierungen dieser Gleichung bei ihrer Anwendung die Totalvariation der numerischen Lösung erhöhen können. Wir beweisen, daß die Oszillationen, die entstehen, in ihrer Größe durch eine zur Zeitschrittweite proportionale Konstante beschränkt bleiben, so daß wir *TV flow* als selbst-stabilisierend bezeichnen können.

Basierend auf Ideen der Bildverarbeitung und der vollständig diskreten Filterkonstruktion entwickeln wir im weiteren Verlauf der Arbeit drei neue Klassen nichtlinearer, diskreter Filter.

Die erste Filterklasse ergibt sich aus den verschiedenen Diskretisierungen einer neuen kontinuierlichen Filtergleichung, die als Extrema-Diffusion interpretiert werden kann. Alle resultierenden Filter sind parameterfrei und besitzen eine sehr einfache Struktur. Bei der Implementierung in Differenzenverfahren höherer Ordnung zeigt sich, daß insbesondere die konservative Diskretisierung der konservativ formulierten Modellgleichung für ein- und zweidimensionale Testfälle zu sehr guten Ergebnissen führt. Die Berechnung der Konvergenzordnung weist nach, daß das resultierende, neue Lösungsverfahren ein Differenzenverfahren höherer Ordnung ist. Darüber hinaus demonstriert die Anwendung des Verfahrens auf eindimensionale Riemann-Probleme mit sonischen Punkten, daß der nichtlineare Filter auch als geeigneter Entropie-Fix eingesetzt werden kann und für das Lax-Wendroff-Verfahren zu besseren Ergebnissen als der klassische Entropie-Fix von Harten führt. Aufgrund der einfachen Filterstruktur ist es gelungen, ein rechentechnisch schnelles und qualitativ gutes Verfahren für das Lösen hyperbolischer Erhaltungsgleichungen zu konstruieren.

Bei der Filterimplementierung in das Zentrale-Differenzen-Verfahren zweiter Ordnung zeigt sich allerdings, daß erst durch eine künstliche Verkleinerung der Zeitschrittweite ein stabiles und konvergentes Lösungsverfahren entsteht.

Da einige andere Diskretisierungen der kontinuierlichen Gleichung zu Verschiebungen der Stoßlage führen, untersuchen wir das Amplituden- und Phasenverhalten eines dieser Filter mit der von Neumannschen Stabilitätsanalyse. Aufschluß über weitere Ursachen dieser Verschiebungsfehler ergeben sich nicht.

Eine zweite Klasse neuer Filtermethoden für das numerische Lösen hyperbolischer Erhaltungsgleichungen konstruieren wir mit Hilfe der Signalzerlegung, die darauf beruht, daß der zu filternde Datensatz vor der Filteranwendung in zwei Teilsignale zerlegt wird. Wir zeigen, daß diese Art der Filterimplementierung qualitativ mit dem Soft Wavelet Shrinkage der Bildverarbeitung übereinstimmt und führen eine Analyse der entstehenden Diffusionsterme durch. Numerische Testfälle zeigen, daß das Filterverfahren die Oszillationen des Lax-Wendroff-

Schemas bis auf eine Restoszillation glättet und Unstetigkeiten erhält. Bei der Filteranwendung auf das Zentrale-Differenzen-Verfahren muß die Zeitschrittweite wieder künstlich verkleinert werden, um Stabilität und Konvergenz zu gewährleisten. Darüber hinaus erlaubt die Kenntnis der Diffusionsterme die direkte Implementierung der Gesamtdiffusion in das Differenzenverfahren, so daß wir ein weiteres, neues Lösungsverfahren erhalten. Da den berechneten Diffusionstermen das Zentrale-Differenzen-Verfahren zugrundeliegt, ist auch hier eine Verkleinerung der hyperbolischen Zeitschrittweite notwendig. Die Berechnung der experimentellen Konvergenzordnung zeigt, daß das resultierende Verfahren ein Verfahren höherer Ordnung ist. Es liefert selbst für Riemann-Probleme mit sonischen Punkten stabile Ergebnisse und kann daher als Entropie-Fix für das Zentrale-Differenzen-Verfahren interpretiert werden, was seine Leistungsfähigkeit beeindruckend dokumentiert.

Die dritte Klasse neuer, nichtlinearer Filterroutinen entwickeln wir, indem wir den zu filternden Datensatz betrachten und durch rein diskrete Operationen Algorithmen herleiten, die Oszillationen vermindern. Vorteil dieses diskreten Vorgehens ist, daß Probleme, die sich bei der Diskretisierung kontinuierlicher Filtergleichungen ergeben, gar nicht erst auftreten können. In Anlehnung an die diskreten Filter von Engquist, Lötstedt und Sjögreen konstruieren wir eine Reihe neuer Filterroutinen, die bei der Implementierung in das Lax-Wendroff- und Zentrale-Differenzen-Verfahren für skalare und vektorielle Testprobleme teilweise zu guten numerischen Resultaten führen, ohne daß Zeitschrittverkleinerungen vorgenommen werden müssen. Insbesondere die invariante zweite Filterversion, die ein Entstehen neuer Extremwerte verhindert, liefert bei Anwendung auf das Lax-Wendroff-Schema ein Verfahren höherer Ordnung, das Oszillationen glättet und Unstetigkeiten in Höhe und Lage erhält. Für Riemann-Probleme mit sonischen Punkten ergibt das Verfahren stabile Lösungen, die einen Dateneinbruch in der Umgebung der Unstetigkeit aufweisen, doch mit Hilfe einer zusätzlichen Hintergrunddiffusion können auch für diese Testprobleme gute numerische Resultate erzielt werden.

Insgesamt zeigt die vorliegende Arbeit, daß einerseits mit Hilfe nichtlinearer Filterroutinen der Bildverarbeitung, andererseits durch vollständig diskret konstruierte Filter, neue nichtlineare Filteralgorithmen entwickelt werden können, die wir als nichtlineare Dissipation interpretieren und so in Differenzenverfahren höherer Ordnung integrieren können, daß sich neue und einfache Methoden für das numerische Lösen hyperbolischer Erhaltungsgleichungen ergeben. Die resultierenden Verfahren sind in ihrer Leistungsfähigkeit teilweise mit technisch aufwendigen, hochauflösenden Methoden vergleichbar, besitzen aber im Gegensatz zu diesen Verfahren einfache und explizit gegebene Diffusionsterme. Es wird allerdings deutlich, daß Filteralgorithmen, die bei der Anwendung auf stationäre, verrauschte Bilder gute Ergebnisse erzielen, bei der Implementierung nach jedem Zeitschritt eines Differenzenverfahrens zu Verschiebungen der Unstetigkeiten in den numerischen Lösungen führen können. Darüber hinaus bereitet das Filtern des Zentrale-Differenzen-Verfahrens zweiter Ordnung Schwierigkeiten. Nichtlineare Filter, die bei der Integration in das Lax-Wendroff-Verfahren gute und stabile Lösungen liefern, erzeugen bei ihrer Anwendung auf das Zentrale-Differenzen-Verfahren Lösungen, in denen für kleiner werdende Gitterweite neue Oszillationen entstehen. Diese Eigenschaft ist auf das instabile Differenzenverfahren zurückzuführen und kann durch eine künstliche Verkleinerung der Zeitschrittweite verhindert werden, wobei der Rechenaufwand des resultierenden Verfahrens allerdings steigt.

Aufgrund der Nichtlinearität der neu konstruierten Filterroutinen ist eine analytische Un-

tersuchung dieser Filter und der resultierenden Lösungsverfahren bisher nicht möglich. Da numerische Untersuchungen, vor allem für Systeme von Erhaltungsgleichungen in mehreren Raumdimensionen, aufwendig sind und die Aussagekraft der Ergebnisse beschränkt ist, sollte die Entwicklung geeigneter mathematischer Analysemethoden, eventuell mit Hilfe der Distributionen- und Maßtheorie, Gegenstand weiterführender Arbeiten sein.

Darüber hinaus ist eine exakte Analyse des Verschiebungsfehlers, der bei der Implementierung einiger Filteralgorithmen entsteht, wünschenswert. Da die Verschiebungsfehler nur auftreten, wenn die Differenzenverfahren mit integrierter Filterroutine auf nichtlineare Testfälle angewendet werden, können diese Effekte nicht nur durch die fehlende Konservativität der Filter begründet werden. Existierende analytische Verfahren, wie die von Neumannsche Stabilitätsanalyse, liefern für diese nichtlineare Problemstellung keine Aussagen.

# Literaturverzeichnis

- [1] L. Alvarez, F. Guichard, P. L. Lions, and J. M. Morel. Axiomatisation et nouveaux opérateurs de la morphologie mathématiques. *Comptes Rendus de l'Académie des Sciences*, t. 315:265–268, 1992. Série I.
- [2] L. Ambrosio. A compactness theorem for a new class of functions of bounded variation. *Bolletino della Unione Matematica Italiana*, VII(4):857–881, 1989.
- [3] D. A. Anderson, J. C. Tannehill, and R. H. Pletcher. *Computational Fluid Mechanics and Heat Transfer*. Series in Computational Methods in Mechanics and Thermal Sciences. Hemisphere Publishing Corporation, 1984.
- [4] F. Andreu, C. Ballester, V. Caselles, and J. M. Mazón. Minimizing total variation flow. *Differential and Integral Equations*, 14(3):321–360, 2001.
- [5] G. Aubert and P. Kornprobst. *Mathematical Problems in Image Processing: Partial Differential Equations and the Calculus of Variations*. Number 147 in Applied Mathematical Sciences. Springer, 2002.
- [6] G. Aubert and L. Vese. A variational method in image recovery. *SIAM J. Num. Anal.*, 34 (5):1984–1979, 1997.
- [7] G. Bellettini, V. Caselles, and M. Novaga. The total variation flow in  $R^N$ . *Journal of Differential Equations*, 184(2):475–525, 2002.
- [8] H. Brezis. *Opérateurs Maximaux Monotones et Semi-Groupes de Contractions dans les Espaces de Hilbert*. North-Holland Publishing Comp., 1973.
- [9] A. Bürgel, M. Breuß, T. Sonar, T. Brox, and J. Weickert. On certain aspects of TV filtering. im Entstehen, 2004.
- [10] F. Catte, P. L. Lions, J. M. Morel, and T. Coll. Image selective smoothing and edge detection by nonlinear diffusion. *SIAM J. Num. Anal.*, 29(1):182–193, 1992.
- [11] T. Cazenave and A. Haraux. *Introduction aux Problèmes d'Evolution Semi-Linéaires*. Mathématiques & Applications. Ellipses, 1990.
- [12] F. Chan, S. Osher, and J. Shen. The digital tv filter and nonlinear denoising. Technical report, Department of Mathematics, UCLA, 1999.
- [13] R. Courant and K. O. Friedrichs. *Supersonic Flow and Shock Waves*. Springer, 1948.

- [14] R. Courant and K. O. Friedrichs und H. Lewy. über die partiellen Differenzengleichungen der mathematischen Physik. *Math. Ann.*, 100:32–74, 1928.
- [15] D. Donoho. De-noising by soft thresholding. *IEEE Transactions on Information Theory*, 41:613–627, 1995.
- [16] B. Enquist, P. Lötstedt, and B. Sjögreen. Nonlinear filters for efficient shock computation. *Mathematics of Computation*, 52 (186):509–537, 1989.
- [17] O. Friedrich. *Gewichtete wesentlich nicht-oszillierende Verfahren auf unstrukturierten Gittern*. PhD thesis, Universität Hamburg, 1999.
- [18] K. Glashoff and H. Kreth. Vorzeichenstabile Differenzenverfahren für parabolische Anfangsrandwertaufgaben. *Numerische Mathematik*, 35:343–354, 1980.
- [19] E. Godlewski and P.-A. Raviart. *Numerical Approximation of Hyperbolic Systems of Conservation Laws*. Number 118 in Applied Mathematical Science. Springer, 1996.
- [20] S. K. Godunov. A difference scheme for numerical computation of discontinuous solution of hydrodynamic equations. *Math. Sbornik*, 47:271–306, 1959.
- [21] T. Grahs. *Nonlinear Anisotropic Diffusion Filters for the Numerical Approximation of Conservation Laws*. PhD thesis, Technische Universität Braunschweig, 2002.
- [22] C. W. Groetsch. *The theory of Tikhonov regularization for Fredholm integral equations of the first kind*. Pitman, Boston, 1984.
- [23] A. Harten. High Resolution Schemes for Hyperbolic Conservation Laws. *Journal of Computational Physics*, 49:357–393, 1983.
- [24] A. Harten. High resolution total-variation-stable finite-difference schemes. *SIAM J. Num. Anal.*, 21:1–23, 1984.
- [25] A. Harten, B. Engquist, S. Osher, and S. Chakravarthy. Uniformly High Order Accurate Essentially Non-Oscillatory Schemes. *Journal of Computational Physics*, 71:231–303, 1987.
- [26] C. Hirsch. *Numerical computation of internal and external flow*, volume 1. John Wiley & Sons, 1988.
- [27] C. Hirsch. *Numerical computation of internal and external flow*, volume 2. John Wiley & Sons, 1990.
- [28] A. Jameson and T. J. Baker. *Solution of the Euler Equations for Complex Configurations*, 1983.
- [29] D. S. Jones. *The theory of generalized functions*. Cambridge University Press, 2nd edition, 1982.
- [30] R. P. Kanwal. *Generalized Functions: Theory and Technique*. Birkhäuser, 2nd edition, 1998.



- [31] N. Kroll and K. Jain. Solution of two-dimensional euler equations – experience with a finite volume code. Technical report, Deutsche Forschungs- und Versuchsanstalt für Luft- und Raumfahrt, Institut für Entwurfaerodynamik, Braunschweig, 1987.
- [32] D. Kröner. *Numerical Schemes for Conservation Laws*. Wiley, Teubner, 1997.
- [33] S. N. Kruzkov. First-order quasilinear equations in several independent variables. *Math. USSR Sbornik*, 10:217–243, 1970.
- [34] P. D. Lax. Shock waves and entropy. In E. A. Zarantello, editor, *Contributions to Nonlinear Functional Analysis*, pages 603–634. Academic Press, 1971.
- [35] P. D. Lax. *Hyperbolic Systems of Conservation Laws and the Mathematical Theory of Shock Waves*. SIAM Publications, 1973.
- [36] P. D. Lax and B. Wendroff. Systems of conservation laws. *Comm. Pure Appl. Math.*, 13:217–237, 1960.
- [37] B. Van Leer. Towards the ultimate conservative difference scheme. i. The quest of monotonicity. *Lecture Notes in Physics*, 18:163–168, 1973.
- [38] B. Van Leer. Towards the ultimate conservative difference scheme. ii. Monotonicity and conservation combined in a second order scheme. *J. Comp. Phys.*, 14:361–370, 1974.
- [39] R. J. LeVeque. *Numerical Methods for Conservation Laws*. Birkhäuser Verlag, 1992.
- [40] A. Madja. *Compressible fluid flow and systems of conservation laws in several space variables*. Springer, 1984.
- [41] J. Málek, J. Nečas, M. Rokyta, and M. Ružička. *Weak and Measure-valued Solutions to Evolutionary PDEs*. Applied Mathematics and Mathematical Computation. Chapman & Hall, 1996.
- [42] A. Marquina and S. Osher. Explicit algorithms for a new time dependent model based on level set motion for nonlinear deblurring and noise removal. Technical report, Department of Mathematics, UCLA, 1999.
- [43] J.-M. Morel and S. Solimini. *Variational Methods in Image Segmentation*. Birkhäuser, 1995.
- [44] J. Von Neumann and R. D. Richtmyer. A method for the numerical calculation of hydrodynamic shocks. *J. Appl. Phys.*, 21:232–237, 1950.
- [45] O. Oleinik. Discontinuous solutions of nonlinear differential equations. *Amer. Math. Soc. Transl. Ser. 2*, 26:95–172, 1957.
- [46] S. Osher and L. I. Rudin. Feature-oriented image enhancement using shock filters. *SIAM J. Num. Anal.*, 27(4):919–940, 1990.
- [47] P. Perona and J. Malik. Scale space and edge detection using anisotropic diffusion. In *Proc. IEEE Comp. Soc.*, pages 12–22, Miami Beach, Nov. 30 - Dec. 2 1987. Workshop on Computer Vision, IEEE Computer Society Press, Washington.

- [48] P. Perona and J. Malik. Scale-space and edge detection using anisotropic diffusion. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12(7):629–639, 1990.
- [49] P. L. Roe. Approximate Riemann solvers, parameter vectors, and difference schemes. *J. Comp. Phys.*, 43:357–372, 1981.
- [50] P. L. Roe and J. Pike. Efficient construction and utilisation of approximate riemann solutions. In R. Glowinski and J. L. Lions, editors, *Computing Methods in Applied Sciences and Engineering*, volume IV, pages 499–518. North-Holland, 1984.
- [51] M. Rosati. Asymptotic behavior of a Geman-Mac Lure discrete model. Technical Report 8, Istituto per le Applicazioni del Calcolo Mauro Picone, 1997.
- [52] L. Rudin and S. Osher. Total variation based image restoration with free local constraints. In *Proceedings of the International Conference on Image Processing*, volume I, pages 31–35, 1994.
- [53] L. Rudin and S. Osher. Total variation based image restoration with free local constraints. *Proc. 1st IEEE ICIP*, 1:31–35, 1994.
- [54] L. Rudin, S. Osher, and E. Fatemi. Nonlinear total variation based noise removal algorithms. *Physica D*, 60:259–268, 1992.
- [55] L. Rudin, S. Osher, and E. Fatemi. Nonlinear total variation based noise removal algorithms. *Physica D*, 60:259–268, 1992.
- [56] L. I. Rudin. *Images, numerical analysis of singularities and shock filters*. PhD thesis, California Institute of Technology, 1987.
- [57] W. Schmidt and A. Jameson. Euler Solvers as an Analysis Tool for Aircraft Aerodynamics. In W. G. Habashi, editor, *Recent Advances in Numerical Methods in Fluids*, volume 4. Pinderidge Press, 1983.
- [58] Yu. I. Shokin. *The Method of Differential Approximation*. Springer, 1983.
- [59] C. W. Shu and C. Hu. Weighted Essentially Non-Oscillatory Schemes on Triangular Meshes. *Journal of Computational Physics*, 150:97–127, 1999.
- [60] J. Smoller. *Shock Waves and Reaction-Diffusion Equations*. Springer, 1983.
- [61] T. Sonar. *Mehrdimensionale ENO-Verfahren: zur Konstruktion nichtoszillatorischer Methoden für hyperbolische Erhaltungsgleichungen*. Advances in Numerical Mathematics. B. G. Teubner Stuttgart, 1997.
- [62] T. Sonar. On the Construction of Essentially Non-Oscillatory Finite Volume Approximations to Hyperbolic Conservation Laws on General Triangulations: Polynomial Recovery, Accuracy and Stencil Selection. *Comput. Methods Appl. Mech. Engrg.*, 140:157–181, 1997.
- [63] G. Steidl and J. Weickert. Relations between soft wavelet shrinkage and total variation denoising. In L. Van Gool, editor, *Pattern Recognition*, volume 2449 of *Lecture Notes in Computer Science*, pages 198–205. Springer, 2002.

- [64] G. Steidl, J. Weickert, T. Brox, P. Mrázek, and M. Welk. On the equivalence of soft wavelet shrinkage, total variation diffusion, total variation regularization. Technical report, Department of Mathematics, University of Bremen, 2003.
- [65] G. Steidl, J. Weickert, T. Brox, P. Mrazek, and M. Welk. On the equivalence of soft wavelet shrinkage, total variation diffusion, total variation regularization, and SIDEs. Technical report, Department of Mathematics, Saarland University, Saarbrücken, Germany, 2003. Accepted for publication in SIAM Journal on Numerical Analysis.
- [66] Bart M. ter Haar Romeny. *Geometric-driven diffusion in computer vision*. Kluwer Academic Publishers, 1994.
- [67] A. N. Tikhonov and V. Y. Arsenin. *Solutions of ill-posed problems*. John Wiley, New York, 1977.
- [68] A. N. Tikhonov and V. Y. Arsenin. *Solutions of Ill-Posed Problems*. Winston and Sons, 1977. Washington, D. C.
- [69] E. F. Toro. *Riemann Solvers and Numerical Methods for Fluid Dynamics*. Springer, 1997.
- [70] L. Vese. *Problèmes variationnels et EDP pour l'analyse d'images et l'évolution de courbes*. PhD thesis, Université de Nice Sophia-Antipolis, November 1996.
- [71] A. I. Vol'pert and S. I. Hudjaev. *Analysis in classes of discontinuous functions and equations of mathematical physics*. Dordrecht, 1985.
- [72] W. Walter. *Einführung in die Theorie der Distributionen*. Bibliographisches Institut AG, Zürich, 1974.
- [73] R. F. Warming and B. J. Hyett. The modified equation approach to the stability and accuracy analysis of finite-difference methods. *J. Comp. Phys.*, 14:159–179, 1974.
- [74] G. W. Wei. Shock capturing by anisotropic oscillation reduction. Technical report, Departemente of Computing Science, National University of Singapore, 2002. preprint, <http://www.math.ntnu.no/conservation/2002/007.html>.
- [75] J. Weickert. Theoretical foundations of anisotropic diffusion in image processing. *Computing Supplement*, 11:221–236, 1996.
- [76] J. Weickert. *Anisotropic Diffusion in Image Processing*. Teubner, 1998.
- [77] P. Wesseling. *Principles of Computational Fluid Dynamics*. Number 29 in Springer Series in Computational Mathematics. Springer, 2001.
- [78] G. Whitham. *Linear and Nonlinear Waves*. Wiley-Interscience, 1974.
- [79] P. Woodward and P. Colella. The numerical simulation of two-dimensional fluid flow with strong shocks. *J. Comput. Phys.*, 54:115–173, 1984.
- [80] S. T. Zalesak. A preliminary comparison of modern shock-capturing schemes: linear advection. In R. Vichnevetsky and R. S. Stepleman, editors, *Advances in Computer Methods for Partial Differential Equations*, volume IV, pages 15–22. IMACS, 1987.



## Zusammenfassung

In der vorliegenden Arbeit entwickeln wir neue nichtlineare, diskrete Filteralgorithmen für das numerische Lösen hyperbolischer Erhaltungsgleichungen. Die Konstruktion der Filteralgorithmen basiert einerseits auf Filterroutinen, die in der Bildverarbeitung entwickelt wurden und zum Wiederherstellen verrauschter Bilder eingesetzt werden, andererseits auf vollständig diskreten Filterkonstruktionen, wie sie erstmals von Engquist, Lötstedt und Sjögreen durchgeführt wurden. Die zugehörigen kontinuierlichen Gleichungen der jeweiligen Filter zeigen, daß beide Ansätze verwandte Filtermethoden liefern. Die konstruierten Filter interpretieren wir als nichtlineare Dissipation und integrieren sie in einfache Differenzenverfahren höherer Ordnung, die in ihren numerischen Lösungen für hyperbolische Erhaltungsgleichungen Oszillationen in der Umgebung von Unstetigkeiten erzeugen. Ziel der Filterintegration in diese Verfahren ist es, die durch das Differenzenverfahren verursachten Oszillationen bei gleichzeitiger Erhaltung der Unstetigkeiten zu glätten und damit neue und einfache Methoden für das numerische Lösen hyperbolischer Erhaltungsgleichungen zu konstruieren.

Zunächst zeigen wir für zwei bekannte TV-Filter der Bildverarbeitung, daß sie auf den Kontext numerischer Approximationen von Erhaltungsgleichungen übertragen und als nichtlineare Dissipation in einfache Differenzenverfahren höherer Ordnung integriert werden können. Eindimensionale skalare beziehungsweise zweidimensionale skalare und vektorielle Testfälle demonstrieren, daß beide Filter bei Anwendung auf stationäre Lösungen von Erhaltungsgleichungen Oszillationen in der Umgebung von Unstetigkeiten reduzieren und dabei Höhe und Lage der Unstetigkeiten weitgehend erhalten. Die Filteranwendung nach jedem Zeitschritt eines Differenzenverfahrens zeigt, daß sie selbst das instabile Zentrale-Differenzen-Verfahren zweiter Ordnung stabilisieren und eine von Oszillationen befreite Lösung erzeugen. Allerdings wird auch deutlich, daß einer der Filter, der TV-Filter von Chan, Osher und Shen, die Lage der Unstetigkeit in der numerischen Lösung gegenüber der tatsächlichen Lage, abhängig von der Wahl der Filterparameter, verschoben darstellt. Nähere Betrachtungen zeigen, daß eine Ursache für diese Verschiebung der Stoßlage die Tatsache ist, daß der diskrete Filter nur für wenige Parameterkombinationen konservativ ist. Obwohl der zweite Filter, ein TV-Filter von Marquina und Osher, auf dem gleichen kontinuierlichen Modell wie der Filter von Chan, Osher und Shen beruht, bewirkt seine Implementierung nach jedem Zeitschritt eines Differenzenverfahrens höherer Ordnung im Wesentlichen keine Verschiebung der Stoßlage. Das unterschiedliche numerische Verhalten der beiden diskreten Filter zeigt, daß Eigenschaften der kontinuierlichen Filtergleichung bei der Diskretisierung verloren gehen können und die Wirkungsweise diskreter Filter stark von der gewählten Diskretisierung abhängt. Aufbauend auf der zugrundeliegenden, kontinuierlichen Filtergleichung entwickeln wir eine Reihe neuer Diskretisierungen und zeigen die Leistungsfähigkeit der resultierenden Filteralgorithmen für eindimensionale skalare Testfälle. Wir erhalten unter anderem einen parame-

terfreien Filter, der als einfache nichtlineare Dissipation das Zentrale-Differenzen-Verfahren stabilisiert, die Oszillationen glättet und Lage und Höhe der Unstetigkeit erhält.

Eine Untersuchung der Gleichung  $u_t = (\operatorname{sgn} u_x)_x$ , die einerseits den neuen Filterroutinen zugrundeliegt, andererseits im Kontext der Bildverarbeitung als *TV flow* bekannt ist und als TV-verringern gilt, zeigt, daß die verschiedenen Diskretisierungen dieser Gleichung bei ihrer Anwendung die Totalvariation der numerischen Lösung erhöhen können. Wir beweisen, daß die Oszillationen, die entstehen, in ihrer Größe durch eine zur Zeitschrittweite proportionale Konstante beschränkt bleiben, so daß wir *TV flow* als selbst-stabilisierend bezeichnen können.

Im weiteren Verlauf der Arbeit entwickeln wir, basierend auf Ideen der Bildverarbeitung und der vollständig diskreten Filterkonstruktion, drei neue Klassen nichtlinearer, diskreter Filter. Die erste Filterklasse ergibt sich aus verschiedenen Diskretisierungen einer neuen kontinuierlichen Filtergleichung, die als Extrema-Diffusion interpretiert werden kann. Alle resultierenden Filter sind parameterfrei und besitzen eine sehr einfache Struktur. Bei der Implementierung in Differenzenverfahren höherer Ordnung zeigt sich, daß insbesondere die konservative Diskretisierung der konservativ formulierten Modellgleichung für ein- und zweidimensionale Testfälle zu sehr guten Ergebnissen führt und das resultierende Verfahren ein Differenzenverfahren höherer Ordnung ist. Darüber hinaus demonstriert die Anwendung des neuen Lösungsverfahrens auf eindimensionale Riemann-Probleme mit sonischen Punkten, daß der nichtlineare Filter auch als geeigneter Entropie-Fix eingesetzt werden kann und für das Lax-Wendroff-Verfahren zu besseren Ergebnissen als der klassische Entropie-Fix von Harten führt.

Bei der Filterimplementierung in das Zentrale-Differenzen-Verfahren zweiter Ordnung zeigt sich, daß erst durch eine künstliche Verkleinerung der Zeitschrittweite ein stabiles und konvergentes Lösungsverfahren entsteht.

Da einige andere Diskretisierungen der kontinuierlichen Gleichung zu Verschiebungen der Stoßlage führen, untersuchen wir das Amplituden- und Phasenverhalten eines dieser Filter mit der von Neumannschen Stabilitätsanalyse. Aufschluß über weitere Ursachen dieser Verschiebungsfehler ergeben sich nicht.

Eine zweite Klasse neuer Filtermethoden für das numerische Lösen hyperbolischer Erhaltungsgleichungen konstruieren wir mit Hilfe der Signalzerlegung, die darauf beruht, daß der zu filternde Datensatz vor der Filteranwendung in zwei Teilsignale zerlegt wird. Wir zeigen, daß diese Art der Filterimplementierung qualitativ mit dem Soft Wavelet Shrinkage der Bildverarbeitung übereinstimmt und führen eine Analyse der entstehenden Diffusionsterme durch. Numerische Testfälle zeigen, daß das Filterverfahren die Oszillationen des Lax-Wendroff-Schemas bis auf eine Restoszillation glättet und Unstetigkeiten erhält. Bei der Filteranwendung auf das Zentrale-Differenzen-Verfahren muß die Zeitschrittweite wieder künstlich verkleinert werden, um Stabilität und Konvergenz zu gewährleisten. Darüber hinaus erlaubt die Kenntnis der Diffusionsterme die direkte Implementierung der Gesamtdiffusion in das Differenzenverfahren, so daß wir ein weiteres, neues Lösungsverfahren erhalten. Da den berechneten Diffusionstermen das Zentrale-Differenzen-Verfahren zugrundeliegt, ist auch hier eine Verkleinerung der hyperbolischen Zeitschrittweite notwendig. Die Berechnung der experimentellen Konvergenzordnung zeigt, daß das resultierende Verfahren ein Verfahren höherer Ordnung ist. Es liefert selbst für Riemann-Probleme mit sonischen Punkten stabile Ergebnisse und kann daher als Entropie-Fix für das Zentrale-Differenzen-Verfahren interpretiert werden, was seine Leistungsfähigkeit beeindruckend dokumentiert.

Die dritte Klasse neuer, nichtlinearer Filterroutinen entwickeln wir, indem wir den zu filternden Datensatz betrachten und durch rein diskrete Operationen Algorithmen herleiten,

die Oszillationen vermindern. Vorteil dieses diskreten Vorgehens ist, daß Probleme, die sich bei der Diskretisierung kontinuierlicher Filtergleichungen ergeben, gar nicht erst auftreten können. In Anlehnung an die diskreten Filter von Engquist, Lötstedt und Sjögreen konstruieren wir eine Reihe neuer Filterroutinen, die bei der Implementierung in das Lax–Wendroff– und Zentrale–Differenzen–Verfahren für skalare und vektorielle Testprobleme teilweise zu guten numerischen Resultaten führen, ohne daß Zeitschrittverkleinerungen vorgenommen werden müssen. Insbesondere die invariante zweite Filterversion, die das Entstehen neuer Extremwerte verhindert, liefert bei Anwendung auf das Lax–Wendroff–Schema ein Verfahren höherer Ordnung, das Oszillationen glättet und Unstetigkeiten in Höhe und Lage erhält. Für Riemann–Probleme mit sonischen Punkten ergibt das Verfahren stabile Lösungen, die mit Hilfe einer zusätzlichen Hintergrunddiffusion weiter verbessert werden können.

Insgesamt zeigt die vorliegende Arbeit, daß einerseits mit Hilfe nichtlinearer Filterroutinen der Bildverarbeitung, andererseits durch vollständig diskret konstruierte Filter, neue nichtlineare Filteralgorithmen entwickelt werden können, die wir als nichtlineare Dissipation interpretieren und so in Differenzenverfahren höherer Ordnung integrieren können, daß sich neue und einfache Methoden für das numerische Lösen hyperbolischer Erhaltungsgleichungen ergeben. Die resultierenden Verfahren sind in ihrer Leistungsfähigkeit teilweise mit technisch aufwendigen, hochauflösenden Methoden vergleichbar, besitzen aber im Gegensatz zu diesen Verfahren einfache und explizit gegebene Diffusionsterme. Es wird allerdings deutlich, daß Filteralgorithmen, die für die Anwendung auf stationäre verrauschte Bilder gute Ergebnisse erzielen, bei der Implementierung nach jedem Zeitschritt eines Differenzenverfahrens zu Verschiebungen der Unstetigkeiten in den numerischen Lösungen führen können. Darüber hinaus bereitet das Filtern des Zentrale–Differenzen–Verfahrens zweiter Ordnung Schwierigkeiten. Nichtlineare Filter, die bei der Integration in das Lax–Wendroff–Verfahren gute und stabile Lösungen liefern, erzeugen bei ihrer Anwendung auf das Zentrale–Differenzen–Verfahren Lösungen, in denen für kleiner werdende Gitterweite neue Oszillationen entstehen. Diese Eigenschaft ist auf das instabile Differenzenverfahren zurückzuführen und kann durch eine künstliche Verkleinerung der Zeitschrittweite verhindert werden, wobei allerdings der Rechenaufwand des resultierenden Verfahrens steigt.